

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

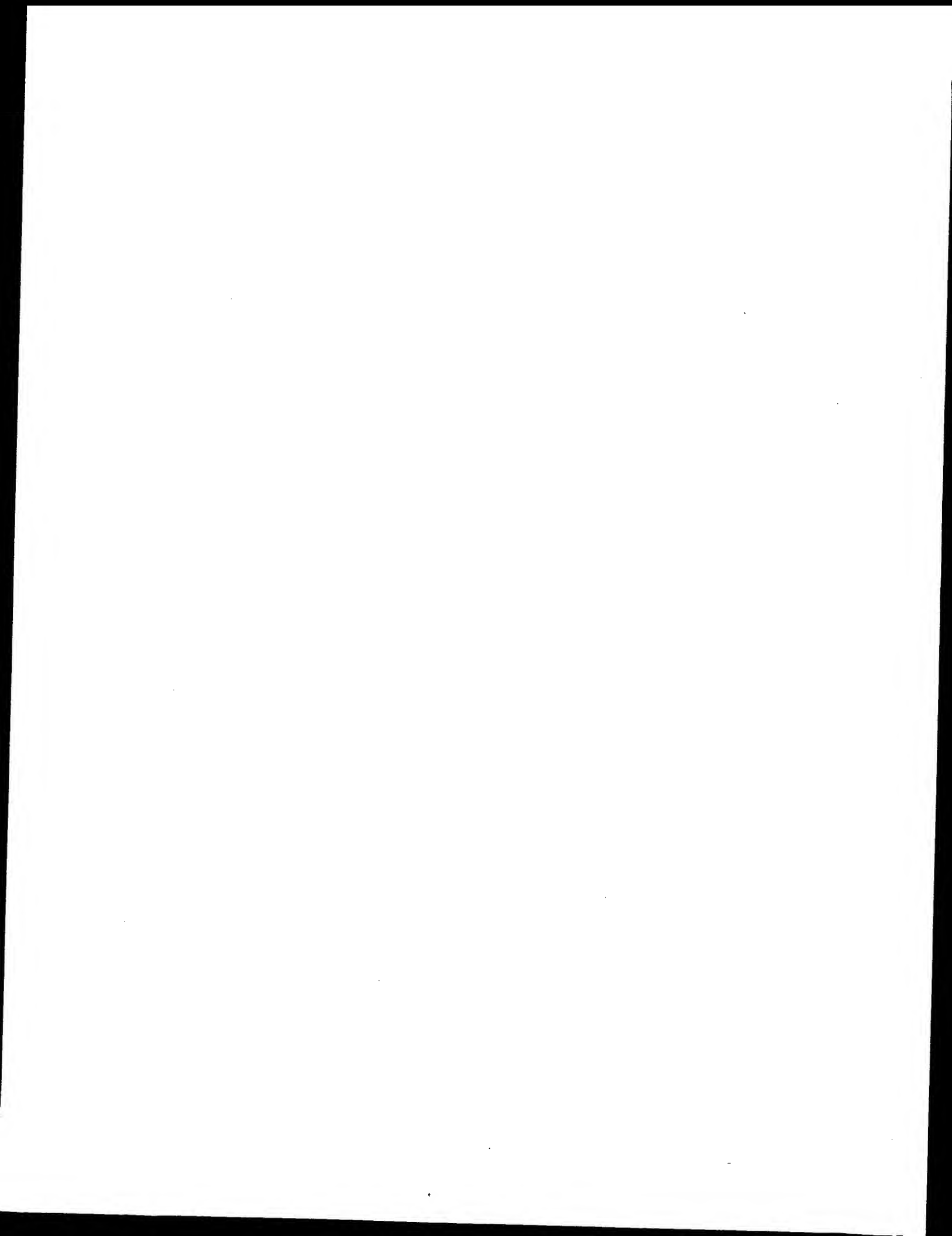
STAFF USE ONLY

Date completed: 08-28-03
Searcher: Beverly C 4994
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

BEST AVAILABLE COPY



Walden
09/937982
Seq. IDs 1-3 w/ inters

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 27, 2003, 00:51:44; Search time 130.55 Seconds
(without alignments)
7529.936 Million cell updates/sec

Title: US-09-937-982-1
Perfect score: 30
Sequence: 1 tccggtccagcgcctctgtttgatgct 30
Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0
Searched: 32997241 seqs, 16383922548 residues
Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Pending patents NA Main:*
- 1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
 - 2: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
 - 3: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
 - 4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
 - 5: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
 - 6: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
 - 7: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
 - 8: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
 - 9: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
 - 10: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
 - 11: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
 - 12: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
 - 13: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
 - 14: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
 - 15: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
 - 16: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
 - 17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
 - 18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
 - 19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
 - 20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
 - 21: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
 - 22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
 - 23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
 - 24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
 - 25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
 - 26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
 - 27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
 - 28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
 - 29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
 - 30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
 - 31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
 - 32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
 - 33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
 - 34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
 - 35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
 - 36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
 - 37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
 - 38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
 - 39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
 - 40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
 - 41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
 - 42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
 - 43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	33	US-09-773-260-1
2	30	100.0	30	39	US-09-937-982-1
3	28	93.3	1133	33	US-09-773-260-3
4	28	93.3	1133	39	US-09-937-982-3

Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli

5 21.6 72.0 1836 53 US-10-437-963-45955
6 21.6 72.0 35740 27 US-09-620-392-50458
7 21.6 72.0 35740 31 US-09-702-134-25459
8 21.6 72.0 35740 34 US-09-815-264-71453
9 20.8 69.3 1694969 91 US-60-361-742-1692
10 20.6 68.7 1287 21 US-09-417-507-19466
11 20.4 68.0 300 37 US-09-881-797-4336
12 20.4 68.0 300 44 US-10-021-698-4336
13 20.4 68.0 599 22 US-09-534-859-763
14 20.4 68.0 599 34 US-09-819-091A-29293
15 20.4 68.0 5015 23 US-09-534-859-763
16 20.4 68.0 5015 34 US-09-803-736-763
17 20.4 68.0 80741 101 US-60-465-241-52917
18 20.4 68.0 80741 101 US-60-466-412-86873
19 20.4 66.7 502 18 US-09-270-849B-59175
20 20.4 66.7 599 28 US-09-654-617-184307
21 20.4 66.7 599 30 US-09-684-016-184307
22 19.8 66.0 573 20 US-09-371-168-4975
23 19.8 66.0 573 34 US-09-801-833-4975
24 19.8 66.0 697 20 US-09-371-168-7927
25 19.8 66.0 697 34 US-09-801-833-7927
26 19.8 66.0 3534 80 US-60-258-273-214
27 19.8 66.0 6802 40 US-09-949-016-2442
28 19.8 66.0 6868 38 US-09-919-039-39
29 19.8 66.0 6868 46 US-10-133-013-139
30 19.8 66.0 6902 28 US-09-654-788A-130
31 19.8 66.0 6902 28 US-09-654-788A-130
32 19.8 66.0 6902 28 US-09-654-788A-130
33 19.8 66.0 6905 1 PCT-US02-21338-77
34 19.8 66.0 6905 1 PCT-US02-21338-77
35 19.8 66.0 6906 47 US-10-188-832-77
36 19.8 66.0 6981 1 PCT-US01-24708-636
37 19.8 66.0 6981 1 PCT-US02-07826-299
38 19.8 66.0 6981 1 PCT-US02-10873A-255
39 19.8 66.0 6981 1 PCT-US02-25766-2609
40 19.8 66.0 6981 1 PCT-US02-30182-977
41 19.8 66.0 6981 2 PCT-US01-24708-636
42 19.8 66.0 6981 36 US-09-873-319-636
43 19.8 66.0 6981 42 US-09-960-706-977
44 19.8 66.0 6981 42 US-09-968-007A-424
45 19.8 66.0 6981 42 US-09-968-007A-424

ALIGNMENTS

RESULT 1
US-09-773-260-1
; Sequence 1, Application US/09773260
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/773,260
; FILING DATE: 31-Jan-2001
; CLASSIFICATION: <Unknown>

Sequence 45955, A
Sequence 50458, A
Sequence 25459, A
Sequence 71453, A
Sequence 1692, Ap
Sequence 19466, A
Sequence 4336, Ap
Sequence 4336, Ap
Sequence 29293, A
Sequence 29293, A
Sequence 763, App
Sequence 763, App
Sequence 52917, A
Sequence 68873, A
Sequence 59175, A
Sequence 184307, A
Sequence 4975, Ap
Sequence 4975, Ap
Sequence 7927, Ap
Sequence 7927, Ap
Sequence 214, App
Sequence 2442, App
Sequence 39, Appl
Sequence 139, App
Sequence 39, Appl
Sequence 130, App
Sequence 130, App
Sequence 130, App
Sequence 77, Appl
Sequence 77, Appl
Sequence 636, App
Sequence 299, App
Sequence 255, App
Sequence 2609, App
Sequence 977, App
Sequence 636, App
Sequence 977, App
Sequence 424, App

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/094,435
FILING DATE: 1998-06-09
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-773-260-1

Query Match 100.0%; Score 30; DB 33; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCCGGATCCAGCGCCTCTGTTTGTATGGCT 30
Db 1 TCCGGATCCAGCGCCTCTGTTTGTATGGCT 30

RESULT 2
US-09-937-982-1
; Sequence 1, Application US/09937982
; GENERAL INFORMATION:
; APPLICANT: DURDEN, Donald L
; TITLE OF INVENTION: DISEASE AND GRAFT VERSUS HOST DISEASE
; FILE REFERENCE: 270/299US
; CURRENT APPLICATION NUMBER: US/09/937,982
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07981
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/127,662
; PRIOR FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Wolinella succinogenes
; US-09-937-982-1

Query Match 100.0%; Score 30; DB 39; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCCGGATCCAGCGCCTCTGTTTGTATGGCT 30
Db 1 TCCGGATCCAGCGCCTCTGTTTGTATGGCT 30

RESULT 3
US-09-773-260-3
; Sequence 3, Application US/09773260
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street

;; Suite 4700
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071-2066
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;;
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: FastSeq for Windows 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/773,260
;; FILING DATE: 31-Jan-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/094,435
;; FILING DATE: 1998-06-09
;; APPLICATION NUMBER: <Unknown>
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 234/274
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510

;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1133 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-773-260-3

Query Match 93.3%; Score 28; DB 39; Length 1133;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CGGATCCAGCGCCTCTGTTTGGTGGCT 30
Db 96 CGGATCCAGCGCCTCTGTTTGGTGGCT 123

RESULT 4
US-09-937-982-3
; Sequence 3, Application US/09937982
; GENERAL INFORMATION:
; APPLICANT: DURDEN, Donald L
; TITLE OF INVENTION: DISEASE AND GRAFT VERSUS HOST DISEASE
; FILE REFERENCE: 270/299US
; CURRENT APPLICATION NUMBER: US/09/937,982
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/US00/07981
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/127,662
; PRIOR FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 1133
; TYPE: DNA
; ORGANISM: Wolinella succinogenes
US-09-937-982-3

Query Match 93.3%; Score 28; DB 39; Length 1133;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CGGATCCAGCGCCTCTGTTTGGTGGCT 30
|||||

Db 96 CGGATCCAGCGCCTCTGTTTGGTGGCT 123

RESULT 5
US-10-437-963-45955
; Sequence 45955, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 45955
; LENGTH: 1836
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4886C.1
US-10-437-963-45955

Query Match 72.0%; Score 21.6; DB 53; Length 1836;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CGGATCCAGCGCCTCTGTTTGGTGGCT 30
Db 1672 CGCTCTGCTCTCTGTTTGGTGGCT 1699

RESULT 6
US-09-620-392-50458
; Sequence 50458, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 50458
; LENGTH: 35740
; TYPE: DNA
; ORGANISM: Oryza sativa
; OTHER INFORMATION: unsure at all n locations
US-09-620-392-50458

Query Match 72.0%; Score 21.6; DB 27; Length 35740;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CGGATCCAGCGCCTCTGTTTGGTGGCT 30
Db 13407 CGCTCTGCTCTCTGTTTGGTGGCT 13434

RESULT 7
US-09-702-134-25459
; Sequence 25459, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 25459
; LENGTH: 35740
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(35740)
; OTHER INFORMATION: unsure at all n locations
US-09-702-134-25459

Query Match 72.0%; Score 21.6; DB 31; Length 35740;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTCTTTGATGGCT 30
Db 13407 CGCTCTGCTCTCTCTCTTTGATGGCT 13434

RESULT 8
US-09-815-264-71453
; Sequence 71453, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR FILING DATE: 2001-03-23
; PRIOR FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 71453
; LENGTH: 35740
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(35740)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-71453

Query Match 72.0%; Score 21.6; DB 34; Length 35740;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTCTTTGATGGCT 30
Db 13407 CGCTCTGCTCTCTCTCTTTGATGGCT 13434

RESULT 9
US-60-361-742-1692
; Sequence 1692, Application US/60361742
; GENERAL INFORMATION:
; APPLICANT: Fidelity Systems, Inc.
```

```
; APPLICANT: Slesarev, Alexei I
; APPLICANT: Mezhevaya, Katja V.
; APPLICANT: Polushin, Nikolai N.
; APPLICANT: Shcherbinina, Olga V.
; APPLICANT: Shakhova, Vera V.
; APPLICANT: Malykh, Andrei G.
; APPLICANT: Kozvavkin, Sergei A.
; TITLE OF INVENTION: Fidelity Systems, Inc.
; FILE REFERENCE: New
; CURRENT APPLICATION NUMBER: US/60/361,742
; CURRENT FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1692
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1692
; LENGTH: 1694969
; TYPE: DNA
; ORGANISM: Methanopyrus kandleri
US-60-361-742-1692

Query Match 69.3%; Score 20.8; DB 91; Length 1694969;
Best Local Similarity 91.7%; Pred. No. 5.6e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCGATCCAGCGCCTCTCTTTGATG 25
Db 1617712 CCGATCCAGCGCCTCTCTTTGATG 1617735

RESULT 10
US-09-417-507-19466/c
; Sequence 19466, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: Keith G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 19466
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: A.fumigatus
US-09-417-507-19466

Query Match 68.7%; Score 20.6; DB 21; Length 1287;
Best Local Similarity 85.2%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCGATCCAGCGCCTCTCTTTGATG 27
Db 1085 TCCGATCCAGCGCCTCTCTTTGATG 1059

RESULT 11
US-09-881-797-4336/c
; Sequence 4336, Application US/09881797
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/09/881,797
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

Thu Aug 28 09:00:59 2003

us-09-937-982-1.rnmpm

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-09-881-797-4336

Query Match      68.0%; Score 20.4; DB 37; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTGATGG 167

RESULT 12
US-10-021-698-4336/c
; Sequence 4336, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-10-021-698-4336

Query Match      68.0%; Score 20.4; DB 44; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTGATGG 167

RESULT 13
US-09-505-532-29293/c
; Sequence 29293, Application US/09505532
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/505,532
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: USSN 60/120,645 1999-02-18; USSN 09/443,025 1999-11-11
; PRIOR FILING DATE: USSN 60/120,645 1999-02-18; USSN 09/443,025 1999-11-12;
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-505-532-29293

Query Match      68.0%; Score 20.4; DB 22; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

RESULT 14
US-09-819-091A-29293/c
; Sequence 29293, Application US/09819091A
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/819,091A
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US 60/108,420
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: US 60/120,645
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 09/443,025
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-819-091A-29293

Query Match      68.0%; Score 20.4; DB 34; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

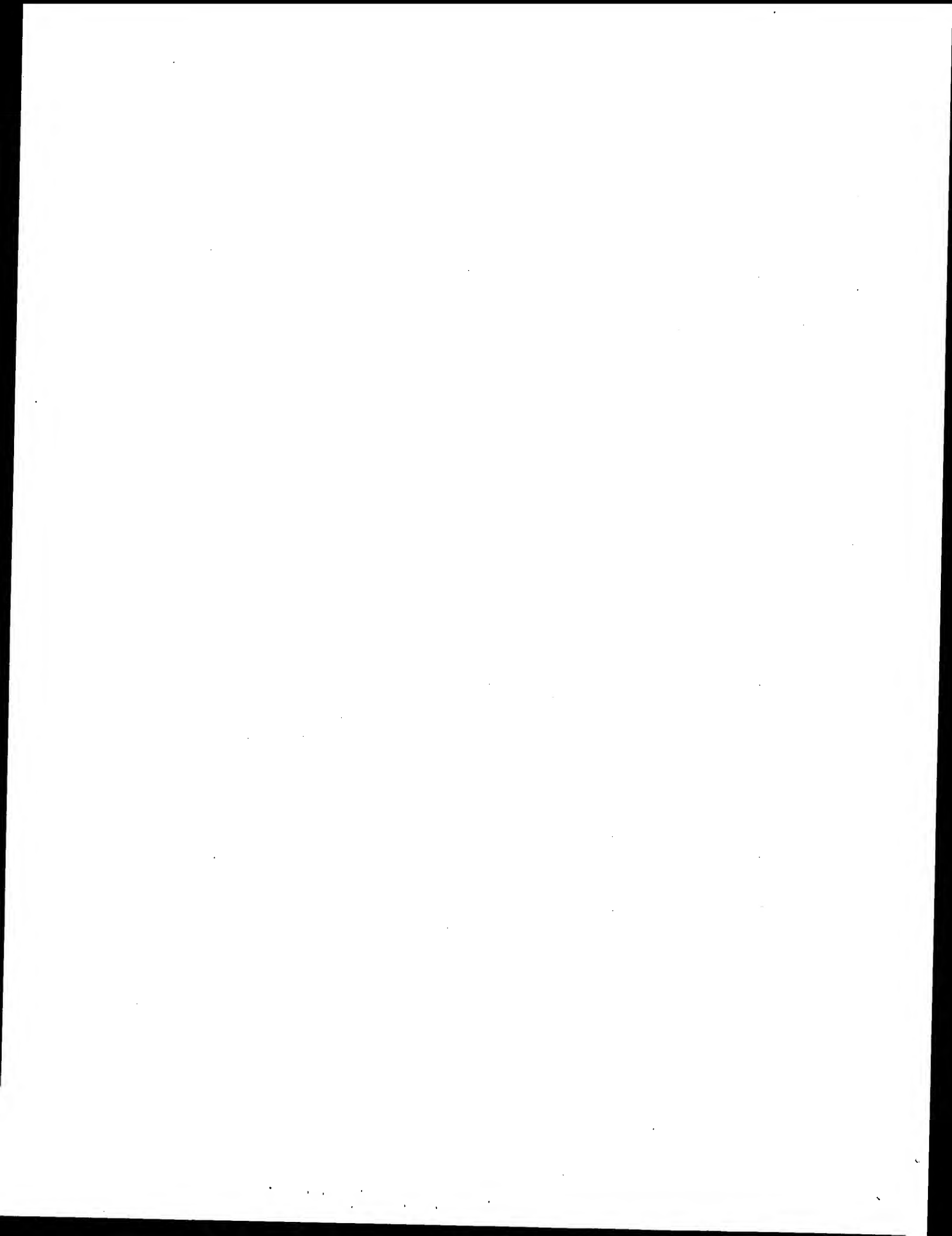
QY 1 TCCGGATCCAGCGCCTCTCTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

RESULT 15
US-09-534-859-763/c
; Sequence 763, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534,859
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1127
; SEQ ID NO 763
; LENGTH: 50015
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-534-859-763

Query Match      68.0%; Score 20.4; DB 23; Length 50015;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTGATGGCT 30
    |||||
Db 15155 TCCGGTTTCATCGCCACTGTTTGTGATCACT 15126

Search completed: August 27, 2003, 10:57:21
Job time : 136.55 secs
```



Thu Aug 28 09:00:59 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 01:11:55 ; Search time 8.8736 Seconds
(without alignments)
3071.976 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30
Sequence: 1 tcggatccagcgctctgtttgatgctt 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 780815 seqs, 454324936 residues

Total number of hits satisfying chosen parameters: 1561630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:
2: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	6 US-10-349-374-1	Sequence 1, Appli
2	28	93.3	1133	6 US-10-349-374-3	Sequence 3, Appli
3	20.6	68.7	2256646	6 US-10-470-565-1	Sequence 1, Appli
4	19.8	66.0	6645	6 US-10-464-368-44	Sequence 44, Appli
5	19.4	64.7	570	6 US-10-470-888-193	Sequence 193, App
6	19.4	64.7	13794	7 US-60-485-450-11872	Sequence 11872, A
7	18.6	62.0	34288	7 US-60-487-610-19242	Sequence 19242, A
8	18.4	61.3	352	6 US-10-629-771-1755	Sequence 1755, Ap
9	18	60.0	2380	6 US-10-302-172-785	Sequence 785, App
10	18	60.0	2452	6 US-10-302-172-786	Sequence 786, App
11	18	60.0	27464	7 US-60-487-610-19431	Sequence 19431, A
12	18	60.0	8059021	5 US-09-947-514-53	Sequence 53, Appli
13	17.8	59.3	545	5 US-09-897-516A-1068	Sequence 1068, Ap
14	17.8	59.3	565	6 US-10-357-930-44477	Sequence 44477, A
15	17.8	59.3	1263	6 US-10-293-244-248	Sequence 248, App
16	17.8	59.3	1386	1 PCT-US02-38594-23	Sequence 23, Appli
17	17.8	59.3	1386	1 PCT-US02-38594-27	Sequence 27, Appli
18	17.8	59.3	1386	1 PCT-US02-38594-29	Sequence 29, Appli
19	17.8	59.3	1386	1 PCT-US02-38594-31	Sequence 31, Appli
20	17.8	59.3	1386	1 PCT-US02-38594-33	Sequence 33, Appli
21	17.8	59.3	2147	6 US-10-357-930-23272	Sequence 23272, A
22	17.8	59.3	2174	6 US-10-357-930-23400	Sequence 23400, A
23	17.8	59.3	2174	6 US-10-357-930-23400	Sequence 23400, A
24	17.8	59.3	2223	1 PCT-US02-38594-37	Sequence 37, Appli
25	17.8	59.3	2223	1 PCT-US02-38594-41	Sequence 41, Appli
26	17.8	59.3	2230	1 PCT-US02-38594-43	Sequence 43, Appli

RESULT 1
US-10-349-374-1
; Sequence 1, Application US/10349374
; GENERAL INFORMATION:
; APPLICANT: Donald L. Durden
; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/349,374
FILING DATE: 21-Jan-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/094,435
FILING DATE: Filed Herewith
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/049,085
FILING DATE: June 9, 1997
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

Sequence 25, Appli
Sequence 3, Appli
Sequence 21, Appli
Sequence 13, Appli
Sequence 39, Appli
Sequence 45, Appli
Sequence 17, Appli
Sequence 15, Appli
Sequence 19, Appli
Sequence 1021, Ap
Sequence 55, Appli
Sequence 54, Appli
Sequence 168, App
Sequence 115, App
Sequence 3539, Ap
Sequence 3540, Ap
Sequence 3541, Ap
Sequence 3542, Ap
Sequence 3543, Ap

ALIGNMENTS

```

;
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-349-374-1

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGGATCCAGCGCCTCTGTTTGTGATGCT 30
Db 1 TCGGATCCAGCGCCTCTGTTTGTGATGCT 30

RESULT 2
US-10-349-374-3
; Sequence 3, Application US/10349374
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/349,374
; FILING DATE: 21-Jan-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,435
; FILING DATE: Filed Herewith
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/049,085
; FILING DATE: June 9, 1997
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-349-374-3

Query Match
Best Local Similarity 93.3%; Score 28; DB 6; Length 1133;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTGATGCT 30
Db 96 CGGATCCAGCGCCTCTGTTTGTGATGCT 123

```

```

RESULT 3
US-10-470-565-1
; Sequence 1, Application US/10470565
; GENERAL INFORMATION:
; APPLICANT: Societe des Produits Nestle S.A.
; TITLE OF INVENTION: NCC2705 - the genome of a Bifidobacterium
; FILE REFERENCE: 80290/WO
; CURRENT APPLICATION NUMBER: US/10/470,565
; CURRENT FILING DATE: 2003-07-29
; PRIOR APPLICATION NUMBER: EP 01102050.0
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 2256646
; TYPE: DNA
; ORGANISM: Bifidobacterium longum
US-10-470-565-1

Query Match
Best Local Similarity 68.7%; Score 20.6; DB 6; Length 2256646;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTGATGCT 29
Db 1702897 CGGATCCAGCGCCTCTGTTTGTGATGCT 1702923

RESULT 4
US-10-464-368-44
; Sequence 44, Application US/10464368
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 44
; LENGTH: 6645
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-464-368-44

Query Match
Best Local Similarity 66.0%; Score 19.8; DB 6; Length 6645;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTGATGCT 28
Db 2434 ATCCAGCGCCTCTGTTTGTGATGCT 2456

RESULT 5
US-10-470-868-193/c
; Sequence 193, Application US/10470868
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 802CIP/PCT
; CURRENT APPLICATION NUMBER: US/10/470,868
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 438
; SEQ ID NO 193
; LENGTH: 570

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (265)...(501)
US-10-470-868-193

Query Match 64.7%; Score 19.4; DB 6; Length 570;
Best Local Similarity 79.3%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 CCGGATCCAGCGCTCTGTTTGGTGGCT 30
DB 36 CTGGATCCTGAGGCTCTTTTGTGGCT 8

RESULT 6
US-60-485-450-11872/c
; Sequence 11872, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 11872
; LENGTH: 13794
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-11872

Query Match 64.7%; Score 19.4; DB 7; Length 13794;
Best Local Similarity 79.3%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 CCGGATCCAGCGCTCTGTTTGGTGGCT 30
DB 6027 CTGGATCCTGAGGCTCTTTTGTGGCT 5999

RESULT 7
US-60-487-610-19242
; Sequence 19242, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 19242
; LENGTH: 34288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-19242

Query Match 62.0%; Score 18.6; DB 7; Length 34288;
Best Local Similarity 84.0%; Pred. No. 90;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 GGATCCAGCGCTCTGTTTGGTGGCT 28
DB 33828 GGATCCAGCGCTGAGTTTGTGGCT 33852

RESULT 8

US-10-629-771-1755/c
; Sequence 1755, Application US/10629771
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: 2300-1598CON
; CURRENT APPLICATION NUMBER: US/10/629,771
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: 09/611,527
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/142,310
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/142,311
; PRIOR FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 3351
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 1755
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-629-771-1755

Query Match 61.3%; Score 18.4; DB 6; Length 352;
Best Local Similarity 78.6%; Pred. No. 73;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGATCCAGCGCTCTGTTTGGTGGCT 28
DB 295 TCCGTTTCCAGGCTCTTGTGTGGTGG 268

RESULT 9

US-10-302-172-785/c
; Sequence 785, Application US/10302172
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Arginine-rich Protein-like Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 ICNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950

OTHER INFORMATION:
US-09-897-516A-1068
Query Match 59.3%; Score 17.8; DB 5; Length 545;
Best Local Similarity 75.9%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TCGGATCCAGCGCCTCTCTTTTGATGGC 29
DB 160 TCGGATCCAGCGCCTCTCTTTTGATGGC 132
RESULT 14
US-10-357-930-44477
; Sequence 44477, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endese, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44477
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-44477
Query Match 59.3%; Score 17.8; DB 6; Length 1263;
Best Local Similarity 75.9%; Pred. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30
DB 616 CTGGATCTGGCGCTTCTACTTTGAGGCT 644

Search completed: August 27, 2003, 11:03:26
Job time : 25.8736 secs

OTHER INFORMATION:
US-09-897-516A-1068
Query Match 59.3%; Score 17.8; DB 5; Length 545;
Best Local Similarity 75.9%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TCGGATCCAGCGCCTCTCTTTTGATGGC 29
DB 160 TCGGATCCAGCGCCTCTCTTTTGATGGC 132
RESULT 14
US-10-357-930-44477
; Sequence 44477, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endese, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44477
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-44477
Query Match 59.3%; Score 17.8; DB 6; Length 565;
Best Local Similarity 75.9%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30
DB 372 CTGGATCTGGCGCTTCTACTTTGAGGCT 400

RESULT 15
US-10-244-248
; Sequence 248, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:51:44 ; Search time 4930.45 Seconds
(without alignments)
7529.936 Million cell updates/sec

Title: US-09-937-982-3
Perfect score: 1133
Sequence: 1 atggcagcagccatcatca.....aagaaagggaatcttctac 1133

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.old.*
3: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US092A COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*

44:	/cgn2_6/ptodata/1/pna/US100A COMB.seq.*
45:	/cgn2_6/ptodata/1/pna/US100B COMB.seq.*
46:	/cgn2_6/ptodata/1/pna/US101A COMB.seq.*
47:	/cgn2_6/ptodata/1/pna/US101B COMB.seq.*
48:	/cgn2_6/ptodata/1/pna/US102A COMB.seq.*
49:	/cgn2_6/ptodata/1/pna/US102B COMB.seq.*
50:	/cgn2_6/ptodata/1/pna/US103A COMB.seq.*
51:	/cgn2_6/ptodata/1/pna/US103B COMB.seq.*
52:	/cgn2_6/ptodata/1/pna/US104A COMB.seq.*
53:	/cgn2_6/ptodata/1/pna/US104B COMB.seq.*
54:	/cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
55:	/cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
56:	/cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
57:	/cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
58:	/cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
59:	/cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
60:	/cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
61:	/cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
62:	/cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
63:	/cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
64:	/cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
65:	/cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
66:	/cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
67:	/cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
68:	/cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
69:	/cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
70:	/cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
71:	/cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
72:	/cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
73:	/cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
74:	/cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
75:	/cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
76:	/cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
77:	/cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
78:	/cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
79:	/cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
80:	/cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
81:	/cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
82:	/cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
83:	/cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
84:	/cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
85:	/cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
86:	/cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
87:	/cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
88:	/cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
89:	/cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
90:	/cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
91:	/cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
92:	/cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
93:	/cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
94:	/cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
95:	/cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
96:	/cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
97:	/cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
98:	/cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
99:	/cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
100:	/cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
101:	/cgn2_6/ptodata/1/pna/US6046 COMB.seq.*
102:	/cgn2_6/ptodata/1/pna/US6047 COMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1133	100.0	1133	39	US-09-937-982-3
2	1129.8	99.7	1133	33	Sequence 3, Appli
C	302.8	26.7	9082	37	Sequence 1293, Ap
C	302.8	26.7	9082	37	US-09-897-516-1294

RESULT 2

US-09-773-260-3

; Sequence 3, Application US/09773260

; GENERAL INFORMATION:

; APPLICANT: Donald L Durden

; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
ASPARAGINASE IN THE TREATMENT OF HUMAN

; HEMATOLOGIC AND AUTOIMMUNE DISEASE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSeq for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/773,260

; FILING DATE: 31-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/094,435

; FILING DATE: 1998-06-09

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 234/274

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1133 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

-US-09-773-260-3

Query Match 99.7%; Score 1129.8; DB 33; Length 1133;
 Best Local Similarity 99.8%; Pred No. 0;
 Matches 1131; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGACAGCCATCATCATCATATAGCAGCGCTGGTGGCGCGGAGCCAT 60
 Db 1 ATGGGACAGCCATCATCATCATATAGCAGCGCTGGTGGCGCGGAGCCAT 60

Qy 61 ATGGCTAGCATGCTGGTGACAGCAATGGTCCGGATCCAGGCGCTCTGTTTGATG 120
 Db 61 ATGGCTAGCATGCTGGTGACAGCAATGGTCCGGATCCAGGCGCTCTGTTTGATG 120

Qy 121 GCTAAACCCCAAGTACTCTAGCCACAGGAGCCATCGCTGGTTCGGGGGAATCT 180
 Db 121 GCTAAACCCCAAGTACTCTAGCCACAGGAGCCATCGCTGGTTCGGGGGAATCT 180

Qy 181 AGCGTCAAGTAGTACTCTGCTGGAGCAGTCAACCGTTGATGATGATGATGATG 240
 Db 181 AGCGTCAAGTAGTACTCTGCTGGAGCAGTCAACCGTTGATGATGATGATGATG 240

Qy 241 CCTGCCATCAAGCAGCCATGACCATCAAGGTGACAGATCTCAAGATTGGTCCCAA 300
 Db 241 CCTGCCATCAAGCAGCCATGACCATCAAGGTGACAGATCTCAAGATTGGTCCCAA 300

Db 241 CCTGCCATCAAGCAGCCATGACCATCAAGGTGAAAGATCTCAAGCATTTGGCTCCCAA 300
 Qy 301 GAGATGACGGTAAAGTGTGGCTTAAACTAGCAAGCGTGTCAATGAGTCTCTCGCCCAA 360
 Db 301 GAGATGACGGTAAAGTGTGGCTTAAACTAGCAAGCGTGTCAATGAGTCTCTCGCCCAA 360
 Qy 361 AAAGAGACCGAAGCGGTGATCATCACCCTAGCACTGACCATGGAAGAGACCGCTTTC 420
 Db 361 AAAGAGACCGAAGCGGTGATCATCACCCTAGCACTGACCATGGAAGAGACCGCTTTC 420
 Qy 421 TTCTCAACCTCACGGTGAAGCAAAACCTGTCTGCTGTAGGGGCCCATGCGTCCA 480
 Db 421 TTCTCAACCTCACGGTGAAGCAAAACCTGTCTGCTGTAGGGGCCCATGCGTCCA 480
 Qy 481 GGCTCTTCCATGAGTGGCGCCCATGATCTCTATAAGCGCGTGAATGAGGATC 540
 Db 481 GGCTCTTCCATGAGTGGCGCCCATGATCTCTATAAGCGCGTGAATGAGGATC 540
 Qy 541 AACAAAGCCTCTACTAAAGAGAGTGGTATTGTGATGAAGATGAGATTCACGCGCC 600
 Db 541 AACAAAGCCTCTACTAAAGAGAGTGGTATTGTGATGAAGATGAGATTCACGCGCC 600
 Qy 601 AGAAGACGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACAGGT 660
 Db 601 AGAAGACGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACAGGT 660
 Qy 661 ABAATCGGCACAGTCTATTATGCAAAAGTGGAGTATTTCACTCAATCGTTCGACCTCAC 720
 Db 661 ABAATCGGCACAGTCTATTATGCAAAAGTGGAGTATTTCACTCAATCGTTCGACCTCAC 720
 Qy 721 ACCCTTGCAGAGTGGTATTGATATTAGCAAAATCGAAGAACTCCCGAGAGTGCATATCTT 780
 Db 721 ACCCTTGCAGAGTGGTATTGATATTAGCAAAATCGAAGAACTCCCGAGAGTGCATATCTT 780
 Qy 781 TAGCTCAACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 Db 781 TAGCTCAACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 Qy 841 GGAATCATCCATCAGGCGATGGGCAATGGGAACTTTCCCTTTGACTCAAAATGCTCTT 900
 Db 841 GGAATCATCCATCAGGCGATGGGCAATGGGAACTTTCCCTTTGACTCAAAATGCTCTT 900
 Qy 901 GAAAAAGCAGCCAAATCAGGCGTAGTCTGCTCGAAGCTCTAGAGTGGGAGTGGTTC 960
 Db 901 GAAAAAGCAGCCAAATCAGGCGTAGTCTGCTCGAAGCTCTAGAGTGGGAGTGGTTC 960
 Qy 961 ACCACCCAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGGCTTACAGAGAGTCTC 1020
 Db 961 ACCACCCAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGGCTTACAGAGAGTCTC 1020
 Qy 1021 AACCTTCAAAAGCCAGAGTCTTCTTATGTTAGCCCTCCACAAACTAGTGTATGATGAG 1080
 Db 1021 AACCTTCAAAAGCCAGAGTCTTCTTATGTTAGCCCTCCACAAACTAGTGTATGATGAG 1080
 Qy 1081 GCGATCCAAAGATCTTCTCCACCTATTATTAATCCAAAGAAAGGAATCTCTTAC 1133
 Db 1081 GCGATCCAAAGATCTTCTCCACCTATTATTAATCCAAAGAAAGGAATCTCTTAC 1133

RESULT 3

US-09-897-516-1293/c

; Sequence 1293, Application US/09897516

; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Huesing, Joseph E.

; APPLICANT: Krasomil-Osterfeld, Karina C.

; APPLICANT: Malvar, Thomas M.

; APPLICANT: Slater, Steven C.

; APPLICANT: Spiridonov, Sergei

; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

; FILE REFERENCE: 38-21(51847)B

QY 367 ACCGAGCGGTGATCATCACCCATGGAACTGACACCATGGGAAGAGACCGCTTTCTTCCTC 426
Db ACTGACGGTTTTGTGTCATCACCCACGGTACAGATACCATGGGAAGAACTGCTTATTTCTTC 5621
QY 427 AACCTCAGGTGAAAGCCAAACCTGCTGCTGCTTTAGGCGCCATGGTCCAGGCTCT 486
Db 5622 GATCTGACCATCTAGTGTCTAATGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCT 5681
QY 487 TCCATGAGTGTGATGGCCCATGAATCTCTATATACGGCGTGAATGATGCGATCAACAA 546
Db 5682 GCTCTGGGTGCTGATGGCCCATGAACCTCTATATGCTGCTGCTGCTGCTGCTGCTGCT 5741
QY 547 GCCTCTACTAACAAGAGTGGTGTGATGATGAACGATGATGATGATGATGATGATGATG 606
Db 5742 GACTCTGCAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5801
QY 607 GCGACCAAGCTCAACACACCGCAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
Db 5802 ATCCGTAACTGAGCACAACCTGAAGTTCAGGCTTTCCAGGCGATGAGCGAGGCCACAA 5861
QY 667 GGCACAGTCTATTATGGCAAGTTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCT 726
Db 5862 GCTCTTCAATCAACGGTAAAGTAACTACTACTCT--GCAGCACCGAGTGAAGCAGAT 5918
QY 727 GCAAGTGAATTTGATATAGCAAAATCGAAGAACTCCCGCAGAGTGTATTTCTTACGCT 786
Db 5919 AAAGCTGTTTTGATGTGAGCAAACTGACTGAACTGCCAAAGTAGTATTTGTATTAAC 5978
QY 787 CACCCGATGATGATGATGTTTTAGTCAATGCGAGCTTTCAGGCGAGGAGCCAAAGAA 846
Db 5979 TACTCTAATGCTTCCGATCTGCGGTTAAAGCAATTTAGAGAAATGATTAAGAGTAT 6038
QY 847 ATCCATCAGGCGATGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906
Db 6039 ATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6098
QY 907 GCAGCCAAATCAGGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 6099 GCGGCTAAAGATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6158
QY 967 CAAGAGCTGAAGTGAATGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Db 6159 CAAATGCTGAAGTGAATGAAGCAAAATACGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 6218
QY 1027 CAAAAAGCCAGAGTCTTCTTATGTTAGCCCTCAACAAATAGTATGATAGAGGGGATC 1086
Db 6219 CAGAAAGCTCGTGTCTTCTGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6278
QY 1087 CAAAAAGCTCTTCTCCACCTATTAAATCCAAAGAA 1118
Db 6279 CAGAAATGTTTTCTAAATATTAAATCTCTGAA 6310

RESULT 8

US-09-897-516-1298
; Sequence 1298, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1298

; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5259)..(6302)
; US-09-897-516-1298

Query Match 26.7%; Score 302.8; DB 37; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAAGTGAATCTCTAGCCACAGAGGAGCACCATCGCTGGTTCGGGGGAATCTAGCGTCT 186
Db 5328 CCAAAACATTAATCTTCTGGCAACGGGTGGCAACATCGCTGGTGGTGAATCTGCAACT 5387
QY 187 AAGAGTAGCTTACTCTGCTGGAGCAGTCAACGTTGATAAGCTTCTTGCAGCGCTCCCTGCC 246
Db 5388 AATCCAGCTACACTGCGGGTCAAGTGGTGTGATTCATTGCTGAATGCTGTCTCTGCT 5447
QY 247 ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306
Db 5448 ATCAAAACATTCCTAATCTGAAGGTGAGCAGGTTCTTAGCATCGGTTCTCAGGACATG 5507
QY 307 ACGGTTAAGTGTGGCTTAAACTAGCCACCGTGTCAATGAGCTCTCGCCCAAAAGAG 366
Db 5508 AACGATCAGGCTCTGGTGAATCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA 5561
QY 367 ACCGAGCGGTGATCATCACCCATGGAACCTGACACCATGGAAGAGACCGCTTTCTTCCTC 426
Db 5562 ACTGACGCTTTGCTGATCACCCACGGTACAGATACCATGGAAGAAATGCTTATTTCTCT 5621
QY 427 AACCTCAGCGTGAAGAACCAAAACCTGCTGCTGCTTAGCGGCATGCGTCCAGGCTCT 486
Db 5622 GATCTGACCACTCAGTGTCTAATCACTATGCTAATGTTGCGCAATGCGTCCATCAAG 5681
QY 487 TCCATGAGTGTGATGGCCCATGAATCTCTATAACCGCGTGAATGATGAGCATCAACAAA 546
Db 5682 GCTCTGGGTGCTGATGGCCCATTGAACCTCTATAATGCTGTTGTTAGCGTCAATATAA 5741
QY 547 GCCTCTACTAACAAGAGTGGTGTGATGATGAACGATGATGATGATGATGATGATGATG 606
Db 5742 GACTCTGCAAAACCGTGGTGTGTTTACTTGTCTGATGAATGATTTCTGTAATTCATGGAG 5801
QY 607 GCAGCAAGCTCAACACACCGCAGTCAATGCTATTTGCTTCCGCCAAACACAGGTAAATC 666
Db 5802 ATCGGTAACTGAGCACAACCTGAAGTTCAGGCTTCCAGGCACTGAAACGAGGCCACAA 5861
QY 667 GCACAGTCTATTATGGCAAGTGGATTTTCACTCAATCCGTTTCGACCTCACACCTTT 726
Db 5862 GGCTTCATTCATAACCGTAAAGTAAACTACTACTCT---GCAGCACCGAGTGAAGCAGAT 5918
QY 727 GCAAGTGAATTTGATATTAAGCAAAATCGAAGAACTCCCGAGTGTGATTTCTTACGCT 786
Db 5919 AAAGCTGTTTTGATGTGCAAAACCTGACTGAACTGCCAAAGTAGGTAATTTGTGTAAT 5978
QY 787 CACCCCGATGATGATGATGTTTTAGTCAATGCGTTTCGACCTCACACCTTTCAAGATC 846
Db 5979 TACTCTAATGCTTCCGATCTGCGGTTAAAGCAATTTAGAGAAATGATTAAGAGTAT 6038
QY 847 ATCCATGAGGCAATGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906
Db 6039 ATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6098
QY 907 GCAGCCAAATCAGGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 6099 GCGGCTAAAGATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6158
QY 967 CAAGAGCTGAAGTGAATGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Db 6159 CAAATGCTGAAGTGAATGAAGCAAAATACGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 6218
QY 1027 CAAAAAGCCAGAGTCTTCTTATGTTAGCCCTCAACAAATAGTATGATAGAGGGGATC 1086
Db 6219 CAGAAAGCTCGTGTCTTCTGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6278
QY 1087 CAAAAAGCTCTTCTCCACCTATTAAATCCAAAGAA 1118
Db 6279 CAGAAATGTTTTCTAAATATTAAATCTCTGAA 6310

Db 6219 CAGAAAGCTCGTGTCTCTGCAATGGCTCTGACTCAGACTGCTGCACACAGCAAAAT 6278
QY 1087 CAAAAGATCTCTCCACCTATTATCAAGAA 1118
Db 6279 CAGAAATGTTTCTAAATATTAATCTCTGAA 6310

RESULT 9

US-60-215-161-1293/c
; Sequence 1293, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1293
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3968)..(5734)
US-60-215-161-1293

Query Match 26.7%; Score 302.8; DB 75; Length 9082;

Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;
QY 127 CCCAAGTACTATCTAGCCAGAGGACCATCGCTGGTTCGGGGAACTCAGCGTC 186
Db 3755 CCAACATTACTGTTCTGGCAACGGGTGGCACATCGTGGTGTGATCTGCACT 3696
QY 187 AGAGTAGTACTCTCTGTGAGCAGTCACCGTTGATGAAGTCTTTGAGCGGTCCTGCC 246
Db 3695 AATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTCGTTGAATGCTGTCTGCT 3636
QY 247 ATCAAGCAGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTCGCTCCCAAGAGATG 306
Db 3635 ATCAAAACATTCGTAATCTGAAGGTGACGAGTTGTTAGCATCGGTCTTCAGGACATG 3576
QY 307 ACGGTAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCTCGGCCCAAAAGAG 366
Db 3575 AACGATCAGTCTGGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAA 3522
QY 367 ACCGAAGCGTGTATCATCACCATGGAATGACACCATGGAAGAGACCGCTTTCTCCTC 426
Db 3521 ACTGACGGTTTGTCTATCACCACCGGTACAGATACCATGGAAGAACTGCTTAATTCCTC 3462
QY 427 AACTCAGCGTGAAGAGCAAAACCTGCTGCTGTAGCGCCATCGCTCCAGGCTCT 486
Db 3461 GATCTGACCACTCAGTGTCTAATACCTATCGTAATGGTTGGCGAATGCGTCCATCAAG 3402
QY 487 TCCATGAGTCTGATGGCCCCCATGAATCTCTATAACCGCGTGAATGATGAGGATCAACAA 546
Db 3401 GCTCTGGTGTCTGATGGCCCATGAACTCTATAATGCTGTGTGTGATGCTCAGATAA 3342
QY 547 GCTCTACTAAACAGAGTGTGATGATGAACATGAGATTCACGCGCCAGAGAA 606
Db 3341 GACTCTGCAACCGTGGTGTGTTTACTTGTGATGAATGATTCGTAATTCATGACGCTGAC 3282
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGAGTAAATC 666
Db 3281 ATCGGTAACTGAGCACTGAAGTTCAGGCTTTCAGGAGTGAACGAGGCCCAAA 3222

QY 667 GGCACAGCTATTATGGCAAGCTCGAGTATTTCACTCAATCCGTTCCAGCTCACACCTT 726
Db 3221 GCTTCATTATCAACGGTAAAGTAACTACTACTCT--GCAGCACCGAGTGAAGCAGAT 3165
QY 727 GCAAGTGAATTTGATATTAGCAAAATCGAAAGTCCCGAGAGTCAATATCTTTAGGCT 786
Db 3164 AAAGCTGTTTTGATGTGAGCAAACTGACTGAACGCAAAAGTAGGTATTTGTATTAAC 3105
QY 787 CACCCGATGATGATGATGTTTTAGTCAATGAGCCCTTCAGCAGGAGCCAAAGGATC 846
Db 3104 TACTCTAATGCTTCGATCTGCCGGTTAAGCATTTGTAGAGAAATGATTATAAAGGTATT 3045
QY 847 ATCCATGAGGATGCGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAATAA 906
Db 3044 ATCAGTGTGTTGTTGTAACGGTAACTATATAAGTCTATCTGAGCTCTCTGTACAA 2985
QY 907 GCAGCAAAATCAGGCTAGTGTGCTGCTGAAAGCTCTAGAGTGGGAGTGGTTCCACACC 966
Db 2984 GCGGCTAAAGATGGTGTGTTGTTGTTGTTCCAAACCGTATTCCTTTCCGTTTCACTACT 2925
QY 967 CAAGAGGCTGAAGTGATGATAAGAACTTGGTTTGTGGCTACAGAGAGTCTCAACCT 1026
Db 2924 CAAATGCTGAAGTTAATGACAGCAAAATACGGTTTGTGCAATCAGAACGCTGAACCCG 2865
QY 1027 CAAAAAGCCAGAGTGTCTTTATGTTAGCCCTCAACAAACTAGTGTATGAGAGGAGATC 1086
Db 2864 CAGAAAGCTGCTGTTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGCACAGCAAAAT 2805
QY 1087 CAAAAGATCTTCTCCACCTATTATTAATCAAGAA 1118
Db 2804 CAGAAATGTTTTTCTAAATATTAATTCCTGAA 2773

RESULT 10

US-60-215-161-1294/c
; Sequence 1294, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1294
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6073)..(6924)
US-60-215-161-1294

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;
QY 127 CCCAAGTACTATCTAGCCAGAGGACCATCGCTGGTTCGGGGAACTCAGCGTC 186
Db 3755 CCAACATTACTGTTCTGGCAACGGGTGGCACATCGTGGTGTGATTCGCAACT 3696
QY 187 AAGAGTAGTACTCTCTGTGAGCAGTCACCGTTGATGAAGTCTTTGAGCGGTCCTGCC 246
Db 3695 AATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTCGTTGAATGCTGTCTGCT 3636
QY 247 ATCAAGCAGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTCGCTCCCAAGAGATG 306

```

Db 3635 ATCAAAAATCTGCTATCTGAAGGTGACGAGTGTGTAGCATCGTCTTCAGGACATG 3576
QY 307 ACGGGTAAGGTGGCTTAACTAGCAACGGTGTCAATGAGTCTTCGCGCCAAAAGAG 366
Db 3575 AAGCATCAGTCTGGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA 3522
QY 367 ACCGAAGCGGTGATCATACCCATGGAACATGACACATGGAAGAGACCGCTTCTTCCTC 426
Db 3521 ACTGACGGTTTGTCTATCACCACCGGTACAGATACCATGAAGAAATCTGTTTCTC 3462
QY 427 AACCTCAGCGTGAAGGACCAAAACCTGCTGCGCTGTAGCGCATCGCTCCAGGCTCT 486
Db 3461 GATCTGACCACTCAGTGTCTATAACCTATCGTAATGTTGGCGAATCGTCCATCAAG 3402
QY 487 TCCATGAGTGTGATGCGCCCATGATCTCTATAAGCCGTGAATGTAGGATCAACAAA 546
Db 3401 GCTCTGGTGTCTGATGGCCCATTAACCTCTATAATGCTGTGTGTAGCGTCAGATAAA 3342
QY 547 GCCTCTACTAACAAAGAGTGGTGTGATGATGAACGATGAGATTCACGCGCCAGAGAA 606
Db 3341 GACTCTGCAACCGTGTGTGTTTACTTGTCTATGATGATTCGTAATTCATGACGTGAC 3282
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGAGTAAATC 666
Db 3281 ATCGGTAACTGAGCAACTGAAGTTCAAGCTTCCAGGAGTGAAGCGGCCACAA 3222
QY 667 GGCACAGTCTATTATGGCAAGTGGCAATGGAAACCTTCCCTTTGACTCAAAATGCTCTT 726
Db 3221 GGCTTCATTAACCGGTAAAGTAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 3165
QY 727 GCAAGTCAGTTCATATATGCAAAATCGAAGACTCCCGAGTGCATATTTTACGCT 786
Db 3164 AAGCTGTTTTGATGTCAGAAACTGACTGAACTGCCAAAGTAGGTATTGTGTATAC 3105
QY 787 CACCCCGATGATCTGATGTTTTAGTCAATGACGCCCTTCAGGAGGAGCCAAAGGAATC 846
Db 3104 TACTCTAATGCTTCCGATCTGCGGTAAAGCAATTTGTAGAGAATGATTATAAGGTATT 3045
QY 847 ATCATGACGAGTGGCAATGGAAACCTTTCCTTCCCTTTGACTCAAAATGCTCTTGAAAA 906
Db 3044 ATCAGTCTGGTGTGTGTAACCGGTAAAGTAAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 2985
QY 907 GCAGCACCAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 2804 CAGGAAATGTTTTCTAAATATTAATCTCTGAA 2773

```

RESULT 11

```

US-60-215-161-1295/c
; Sequence 1295, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

```

; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1295
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6985)..(9081)
; US-60-215-161-1295

```

```

Query Match      26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCAAGTGACTATCTTAGCCACAGAGGACACATCGTGTTCGGGGGAATCTAGCGTC 186
Db 3755 CCAACATTTACTTCTGGCAACGGGTGCAAAATCGCTGGTGGTGAATCTGCAACT 3696
QY 187 AAGAGTAGTACTCTCTGGAGCAGTCACCGTTGATAAGCTTCTGCAGCGCTCCCTGCC 246
Db 3695 AAATCCAGCTACACTGCGGGTCAAGTCCGTGTGATTCATTTGCTGAATGCTGTCTCT 3636
QY 247 ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306
Db 3635 ATCAAAACATTTGCTTAATCTGAAGGTGAGCAGGTTGTTAGCATCGTCTCAGGACATG 3576
QY 307 ACGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATGAGCTCTCCGCCCAAAAGAG 366
Db 3575 AACGATCAGGTGTGGCTGACTCTGCGGAAAAAATCAATG-----CAGATTGTGATAAA 3522
QY 367 ACCGAAGCGGTGATCATCAACCATGGAACCTGACACCATGGAAGAGACCGCTTCTTCCTC 426
Db 3521 ACTGACGGTTTTGTTGATCACCACCGTACAGATACCATGGAAGAACTGCTTATTTCTC 3462
QY 427 AACCTCAGCGTGAAGGACCAAAACCTGCTGTGCTTTGTAGGCGCCATCGCTCCAGGCTCT 486
Db 3461 GATCTGACCACTCAGTGTCTATAACCTATCGTAATGTTGGCGCAATCGTCCATCAACG 3402
QY 487 TCCATGAGTGTGATGCGCCCATGATCTCTATAAGCCGTGAATGTAGGATCAACAAA 546
Db 3401 GCTCTGGTGTCTGATGGCCCATTAACCTCTATAATGCTGTGTGTAGCGTCAGATAAA 3342
QY 547 GCCTCTACTAAACAAAGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 606
Db 3341 GACTCTGCAACCGTGTGTGTTTACTTGTCTATGATGATGATGATGATGATGATGATGATGATGAT 3282
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGAGTAAATC 666
Db 3281 ATCGGTAACTGAGCAACTGAAGTTCAAGCTTCCAGGAGTGAAGCGGCCACAA 3222
QY 667 GGCACAGTCTATTATGGCAAGTGGCAATGGAAACCTTCCCTTTGACTCAAAATGCTCTT 726
Db 3221 GGCTTCATTAACCGGTAAAGTAAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 3165
QY 727 GCAAGTCAGTTCATATATGCAAAATCGAAGACTCCCGAGTGCATATTTTACGCT 786
Db 3164 AAGCTGTTTTGATGTCAGAAACTGACTGAACTGCCAAAGTAGGTATTGTGTATAC 3105
QY 787 CACCCCGATGATCTGATGTTTTAGTCAATGACGCCCTTCAGGAGGAGCCAAAGGAATC 846
Db 3104 TACTCTAATGCTTCCGATCTGCGGTAAAGCAATTTGTAGAGAATGATTATAAGGTATT 3045
QY 847 ATCATGACGAGTGGCAATGGAAACCTTTCCTTCCCTTTGACTCAAAATGCTCTTGAAAA 906
Db 3044 ATCAGTCTGGTGTGTGTAACCGGTAAAGTAAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 2985
QY 907 GCAGCACCAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 2984 CAGGAAATGTTTTCTAAATATTAATCTCTGAA 2773

```


187 AGAGTAGTACTCTGCTGGAGCAGTCAAGTTGATAAGCTTCTTGAGCGCTCCCTGCC 246
5388 AATCCAGCTACACTCGGGTCAAGTCGGTGTGATTCATTGCTGATGCTGTTCTGCT 5447
247 ATCAACGACCTAGCACCACATCAAGGGTGAACAGATCTCAAGATTTGGCTCCCAAGAGATG 306
5448 ATCAAAACATTGCTAATCTGAAAGGTGAGCAGGTTGTTAGCATCGGTTCTCAGGACATG 5507
307 ACGGGTAAGGTGGCTTAACTAGCAAGCGTGTCAATGAGCTCTCGCCCAAAAGAG 366
5508 AACGATCAGGTGGCTGACTCTGGCGAAATAAATCAATG-----CAGATTGTGATAA 5561
367 ACCGAGCGGTGATCATCACCATGGAACCTGACACCATGGAAGAGACCGCTTCTTCTCCT 426
5562 ACTGACCGTGTGTCATCACCCAGGTGACATACCATGGAAGAACTGCTTATTTCTCCT 5621
427 AACCTCAGCGTGAAGCAAAACCTGCTGCTGTTAGCGGCATCGCTCGAGGCTCT 486
5622 GATCTGACCACTCAGTGTGATCAAAACCTATCGTATGTTGGCGAATGCTTCATCAAG 5681
487 TCCATGAGTGTGATGAGCGCCCATGAATCTTATACGCGGTGAATGTAGCGATCAACAAA 546
5682 GCTCTGGGTGCTGATGCGCCCATGAACTCTTATATGCTGTTGTTAGCGTCAAGATAA 5741
547 GCCTCTACTAACAAGAGAGTGTGATGTAAGAGATGAGATTCACGCGCCAGAGAA 606
5742 GACTCTGCAACCGGTGTTGTTTACTTGTCTATGAATGATCTGTAAATTCATGGACGTGAC 5801
607 GGGACCAAGCTCAACACACCGCAGTCAATGCTTGTTCGCCCAACACAGAGTAAAAATC 666
5802 ATCGGTAACTCAGACCACTCAAGTTCAGGCTTCCAGGCACTGAACGCGCCCAACAA 5861
667 GGCACAGTCTATATGCGCAAGTCGAGTATTTCACTCAATCCGTTGCGACCTCACACCTT 726
5862 GCTTCAATCAACGTTAAGTAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 5918
727 GCAAGTGTGATGATTAAGCAAAATCGAAGACTCCCGAGAGTCCGATATCTTTACGCT 786
5919 AAGCTGTGTTTGTGATCTGCGGTTTAAAGCATTTGTAGAGATGATTTATAAGGTATT 6038
787 CACCCGATGATCTGATGTTTGTAGTCAATGACGCGCTTCAAGCAGGAGCAAGGAATC 846
5979 TACTCTAATGCTTCGATCTGCGGTTTAAAGCATTTGTAGAGATGATTTATAAGGTATT 6038
847 ATCCATGAGGATGAGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGAATAA 906
6039 ATCAGTGTGTTGTTGTTAAGCAATCAATCTATAAGTCTATTCTGGACTCTTGTATTA 6098
907 GCAGCAGGATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 966
6099 GCGGCTAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
967 CAAGAGGCTGAAGTGAATGAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1026
6159 CAAATGCTGAAGTGAATGAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6218
1027 CAAAGAGGAGGAGTGTCTTATGTTAGCGCTTCAACAAATAGTGTGATGAGAGCGGATC 1086
6219 CAGAAAGCTGCTGTTCTTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6278
1087 CAAAGATGCTTCTCCACCTTATTAATCCAGAA 1118
6279 CAGGAATGTTTCTTAATATTAATTCCTGAA 6310

RESULT 14
US-60-215-161-1298
; Sequence 1298, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.

APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)A
CURRENT APPLICATION NUMBER: US/60/215,161
CURRENT FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 1298
LENGTH: 9082
TYPE: DNA
ORGANISM: Xenorhabdus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (5259) .. (6302)
US-60-215-161-1298

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAAGTGAATCTCTAGCCACAGAGGACCACTCGCTGTTGGGGGAATCTAGCGTC 186
Db 5328 CCAACATTACTGTTCTGGCAACGGTGGCAATCGCTGTTGGTGAATCTGCAACT 5387
QY 187 AAGAGTAGTACTCTGCTGGAGCAGTCAACCGTTGATAGCTTCTTGCAGCGCTCCCTGCC 246
Db 5388 AATCCAGCTACACTGCGGTCAAGTCGGTGTGATTCATTGCTGAAATGCTGTTCTGCT 5447
QY 247 ATCAACGACCTAGCACCACCTCAAGGTTGAACAGATCTCAAGATTTGGTCCCAAGAGATG 306
Db 5448 ATCAAAACATTGCTTAACTGAAAGTGAAGGTTGTTAGCATCGGTTCTCAGGACATG 5507
QY 307 ACGGTTAAGTGTGGCTTAACTAGCAGCGTGTCAATGAGCTCTCGCCCAAAAGAG 366
Db 5508 AACGATCAGGTGGCTGACTCTGGCGAAATAAATCAATG-----CAGATTGTGATAA 5561
QY 367 ACCGAGCGGTGATCATCACCATGGAACCTGACACCATGGAAGAGACCGCTTCTTCTCCT 426
Db 5562 ACTGACCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5621
QY 427 AACCTCAGCGTGAAGCAAAACCTGCTGCTTGTAGGCGCCATCGCTCAGGCTCT 486
Db 5622 GATCTGACCACTCAGTGTGATTAACCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5681
QY 487 TCCATGAGTGTGATGCGCCCATGAAATCTTATAACGCGGTGAATGTAGGATCAACAAA 546
Db 5682 GCTCTGGGTGCTGATGCGCCCATTTGAACCTCTATAATGCTGTTGTTGTTGTTGTTGTTGTT 5741
QY 547 GCCTCTACTAACAAGAGAGTGTGATGTAAGAGATGAGATTCACGCGCCAGAGAA 606
Db 5742 GACTCTGCAACCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5801
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 666
Db 5802 ATCGGTAACTGAGCAAACTGAAGTTTCAGGCTTCCAGGCACTGAACGCGCCCAACAA 5861
QY 667 GGCACAGTCTATATGCGCAAGTCGAGTATTTCACTCAATCCGTTGCGACCTCACACCTT 726
Db 5862 GCTTCAATCAACGTTAAGTAACTACTACTCT---GCAGCACCAGTGAAGACAGAT 5918
QY 727 GCAAGTGTGATGATGTTTGTAGTCAATGACGCGCTTCAAGCAGGAGCAAGGAATC 846
Db 5919 AAGCTGTGTTTGTGATCTGCGGTTTAAAGCATTTGTAGAGATGATTTATAAGGTATT 6038
QY 787 CACCCGATGATCTGATGTTTGTAGTCAATGACGCGCTTCAAGCAGGAGCAAGGAATC 846
Db 5979 TACTCTAATGCTTCGATCTGCGGTTTAAAGCATTTGTAGAGATGATTTATAAGGTATT 6038
QY 847 ATCCATGAGGATGAGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGAATAA 906
Db 6039 ATCAGTGTGTTGTTGTTAAGCAATCAATCTATAAGTCTATTCTGGACTCTTGTATTA 6098

Sequence 163, App
Sequence 19647, A
Sequence 24, Appl
Sequence 16, Appl
Sequence 31, Appl
Sequence 30, Appl
Sequence 7, Appl
Sequence 1386, Ap
Sequence 231, App
Sequence 52, Appl
Sequence 3945, Ap
Sequence 19587, A
Sequence 13856, A
Sequence 34797, A
Sequence 43649, A
Sequence 87, Appl
Sequence 4130, Ap
Sequence 19352, A

US-09-655-378A-163
US-60-487-610-19647
US-60-494-568-24
US-60-494-568-16
US-09-655-378A-31
US-09-655-378A-30
US-10-426-124-7
US-10-374-780A-1386
US-10-612-783-231
US-10-360-123A-53
US-10-273-573-3945
US-60-487-610-19587
US-10-357-930-13656
US-10-357-930-34797
US-10-357-930-43649
US-10-623-813-87
US-10-603-113-4130
US-60-487-610-19352

38 3.4 144 5
36 3.2 142005 7
35.4 3.1 20256 7
35.4 3.1 164051 7
35 3.1 948 5
32 3.1 1588 6
33 3.0 1637 6
34 3.0 3071 6
35 3.0 68 6
36 3.0 76 6
37 3.0 1239 6
38 3.0 214354 7
39 3.0 384 6
40 3.4 2.9 405 6
41 3.4 2.9 405 6
42 3.4 2.9 56 6
43 3.2 2.9 774 6
44 3.2 2.9 91217 7
45 3.2 2.9 91217 7

ALIGNMENTS

RESULT 1
US-10-349-374-3
; Sequence 3, Application US/10349374
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/349,374
FILING DATE: 21-Jan-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/094,435
FILING DATE: Filed Herewith
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/049,085
FILING DATE: June 9, 1997
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1133 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

27 38 3.4 144 5
28 36 3.2 142005 7
29 35.4 3.1 20256 7
30 35.4 3.1 164051 7
31 35 3.1 948 5
32 32 3.1 1588 6
33 33 3.0 1637 6
34 34 3.0 3071 6
35 35 3.0 68 6
36 36 3.0 76 6
37 37 3.0 1239 6
38 38 3.0 214354 7
39 39 3.0 384 6
40 40 3.4 2.9 405 6
41 41 3.4 2.9 405 6
42 42 3.4 2.9 56 6
43 43 3.2 2.9 774 6
44 44 3.2 2.9 91217 7
45 45 3.2 2.9 91217 7

OM nucleic - nucleic search, using sw model
Run on: August 27, 2003, 01:11:55 ; Search time 335.126 Seconds
(without alignments)
3071.976 Million cell updates/sec

Title: US-09-937-982-3
Perfect score: 1133
Sequence: 1 atgggcagcagccatcatca.....aagaaagggaattctttcac 1133

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 780815 seqs, 454324936 residues

Total number of hits satisfying chosen parameters: 1561630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/prodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	6	US-10-349-374-3
2	302.8	26.7	9082	5	US-09-897-516A-1293
3	302.8	26.7	9082	5	US-09-897-516A-1294
4	302.8	26.7	9082	5	US-09-897-516A-1295
5	302.8	26.7	9082	5	US-09-897-516A-1296
6	302.8	26.7	9082	5	US-09-897-516A-1297
7	302.8	26.7	9082	5	US-09-897-516A-1298
8	296.4	26.2	1830121	6	US-10-329-670-1
9	277.2	24.5	1041	6	US-10-603-114-1219
10	84.4	7.4	966	6	US-10-617-320-791
11	67.4	5.9	82	6	US-10-406-676-3
12	67.4	5.9	1067	1	PCT-US03-06661A-9
13	67.4	5.9	1074	1	PCT-US03-06661A-11
14	65.2	5.8	759	6	US-10-423-156-60
15	63.4	5.6	759	6	US-10-423-156-59
16	61.4	5.4	88	6	US-10-375-209A-40
17	45	4.0	5347	1	PCT-US02-34768-1
18	45	4.0	5347	1	PCT-US02-34769-17
19	44.4	3.9	1365	6	US-10-372-473-8
20	44.2	3.9	1386	6	US-10-281-092-3
21	43.8	3.9	2076	6	US-10-426-124-5
22	42.8	3.8	1371	6	US-10-281-092-7
23	38.4	3.4	740	6	US-10-311-473-9
24	38.4	3.4	1157	6	US-10-311-473-8
25	38.4	3.4	5309	6	US-10-622-220-14
26	38.4	3.4	5537	6	US-10-622-220-13

Db	Seq ID	Seq	Score	DB 5	Length	9082
Db	961	ACCAACCAAGAGCTGAAGTGGATGATAGAAACTTGGTTTTGGGTACAGAGAGTCTC	1020			
Qy	1021	AACCTCAAAAAGCCAGAGTGTCTTTATGTTAGCCCTACCAAAAATAGTGATAGAGAG	1080			
Db	1021	AACCTCAAAAAGCCAGAGTGTCTTTATGTTAGCCCTACCAAAAATAGTGATAGAGAG	1080			
Qy	1081	CGATCCAAAAGATCTTCTCCACCTATTATCCAGAAAGGGAATCTCTTCAC	1133			
Db	1081	CGATCCAAAAGATCTTCTCCACCTATTATCCAGAAAGGGAATCTCTTCAC	1133			
RESULT 2						
US-09-897-516A-1293/c						
; Sequence 1293, Application US/09897516A						
; GENERAL INFORMATION:						
; APPLICANT: Corbin, David R.						
; APPLICANT: Goldman, Barry S.						
; APPLICANT: Hinkle, Gregory J.						
; APPLICANT: Huesing, Joseph E.						
; APPLICANT: Malvar, Thomas M.						
; APPLICANT: Krasomil-Osterfeld, Karina C.						
; APPLICANT: Slater, Steven C.						
; APPLICANT: Spiridonov, Sergei						
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof						
; FILE REFERENCE: 38-21(51847)B						
; CURRENT APPLICATION NUMBER: US/09/897,516A						
; PRIOR FILING DATE: 2001-06-29						
; PRIOR APPLICATION NUMBER: US 60/215,161						
; PRIOR FILING DATE: 2000-06-30						
; NUMBER OF SEQ ID NOS: 8415						
; SEQ ID NO 1293						
; LENGTH: 9082						
; TYPE: DNA						
; ORGANISM: Xenorhabdus sp.						
; FEATURE:						
; NAME/KEY: CDS						
; LOCATION: (3968)..(5734)						
; OTHER INFORMATION:						
US-09-897-516A-1293						
Query Match 26.7%; Score 302.8; DB 5; Length 9082;						
Best Local Similarity 58.1%; Pred No. 2.2e-91;						
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;						
Qy	127	CCCCAAGTGAATCTCTAGCACAGGAGGCACCATCGCTGGTTCGGGGAAATCTAGCGTC	186			
Db	3755	CCAAACATTACTGTTCTGGCAACGGGTGGCACAAATCGCTGGTGGTGGTGAATCTGCAACT	3696			
Qy	187	AAGAGTAGTACTCTCTGGAGCAGTCAACGGTTGATAGCTTCTTGACGCCGCTCCCTGCC	246			
Db	3695	AAATCAGCTACACTCGGGTCAAGTCGGTGTGAATTCATTGCTGAATGCTGTTCTGCTGCT	3636			
Qy	247	ATCAAGACCTAGCCACCACCATCAAGGGTGAACAGATCTCAAGCATTGGCTCCCAAGAGATG	306			
Db	3635	ATCAAAAACATTGCTAATCTGAAAGGTGACGAGTGTGATGACATCGGTTCTCAGACATG	3576			
Qy	307	ACGGGTAAAGTGTGGCTTAACTAGCCACGGTGTCAATGAGTCTCTCGCCCAAAAAGAG	366			
Db	3575	AACGATCAGGTCTGGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA	3522			
Qy	367	ACCGAAGCCGTGATCATCAACCATGAACTGACACCATGGAAGAGACGGCTTCTTCCTC	426			
Db	3521	ACTGACGGTTTTGTCTATCACCACCGTACAGATACCATGAGAGAACTGCTATTTCCTC	3462			
Qy	427	AACCTCAGGTGAAAAGCCAAAACCTGTCTGTCCTGTAGCGCCATCGGTCCAGGCTCT	486			
Db	3461	GATCTGACCACTCAGTGTCTATAAACCTATCGTAATGGTTGGCGCAATGCGTCCATCAACG	3402			
Qy	487	TCCATGAGTGTGATGGCCCCCATGAATCTCTATACCGCCGTGAATGATGCGATCAACAA	546			
Db	3401	GCTCTGGGTGCTGATGGCCCATTAACCTCTATATGCTGTGTTGAGCGTGAATATAA	3342			

547 GCCTCTACTAAACAAAGGAGTGGTATTTGATGAAACGATGAGATTCACGCGCCAGAGAA 606
 3341 GACTCTGCAACCGTGGTGTCTTACTGTCTATGAAATGATCTGTAATTCATGACGTGAC 3282
 607 GCGACCAAGCTCAACACCCAGTCAATGCAATTTGCTTTCGCCCAACACAGGTAAATC 666
 3281 ATCCGTAAGTGAACCACTGAAGTTCAGGCTTTCAGGCGAGTGAACGCGAGGCCACAA 3222
 667 GGCACAGTCTTATATGCAAGTGCAGTATTTTCACTCAATCCGTTCCACCTCACCCTT 726
 3221 GGCCTCAATTCATAACCGTAAAGTAACTACTCTCT---GCAGCACCAGTGAAGCAGAT 3165
 727 GCAAGTGCAGTTGATATTAAGCAAAATCGAAGAACTCCCGAGTGCATATTTCTTACGCT 786
 3164 AAAGCTGTTTGTATGTCAGCAAACTGACTGAAGTGCCTCAAAAGTAGGTATTTGTATAAC 3105
 787 CACCCCGATGATCTGATGTTTATGTAATGCAATGCAAGCCCTTCAGGCGAGGCCAAAGGAATC 846
 3104 TACTCTAATGCTTCCGATCTGCGGTTTAAAGCATTTGTAGAGATGATTATTAAGGTATT 3045
 847 ATCCATGCGAGTGGCAATGGAACCTTTCCTTTGACTCAAAATGCTCTTGAAGAA 906
 3044 ATCAGTGCCTGTTGTTGTAACCGTAACTCTATTAAGTCTATTCTGGACTCTCTGTACAA 2985
 907 GCAGCCAAATCAGCGTAGTCTGCTGCTGCAAGCTCTAGAGTGGCGAGTGGTTCACCAACC 966
 2984 GCGCTTAAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925
 967 CAAGCGCTGAAGTGGATGATGAAGAACTTGGTTTGGTTCAGAGAGTCTCAACCTT 1026
 2924 CAAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2865
 1027 CAAAGGCGAGTGGTCTTCTTATGTTAGCCCTCACCAGTCTGATGATGAGAGGCGATC 1086
 2864 CAGAAAGCTGCTGTTCTTCTGCAATGGCTCTGACTCAGACTGCTGACACAGCAAAAT 2805
 1087 CAAAGATCTTCTCCACCTATTATCAAGAA 1118
 2804 CAGAAATGTTTCTAAATATTATCTCTGAA 2773

RESULT 3

US-09-897-516A-1294/c
 ; Sequence 1294, Application US/09897516A
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Huesing, Joseph E.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spiridonov, Sergei
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)B
 ; CURRENT APPLICATION NUMBER: US/09/897,516A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 8415
 ; SEQ ID NO 1294
 ; LENGTH: 9082
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (6073)..(6924)
 ; OTHER INFORMATION:
 ; US-09-897-516A-1294

Query Match 26.7%; Score 302.8; DB 5; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 2.2e-91;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

127 CCCAAGTGAATCTTCTAGCCACAGGAGGACCATCTGCTGTTTCGGGGGAATCTAGCGTC 186
 3755 CCAAACTTACTGTTCTTGGCAACGGTGGCAAACTGCTGGTGGTGAATCTGCAACT 3696
 187 AAGAGTACTACTCTGCTGGAGCAGTCAACGTTGATAGCTTCTTGCAGCGCTCCCTGCC 246
 3695 AAATCCAGCTACACTGCGGGTCAAGTGGTGTGTTTCAATGCTGAATGCTGTTCTGCT 3636
 247 ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306
 3635 ATCAAAACATTTGTAATCTGAAAGTGAGAGTGTGTAGCATCGTCTCAGGACATG 3576
 307 ACGGTAAAGTGTGGCTTAAACTAGCCAGCGTGTCAATGAGCTCTCGCCCCAAAAGAG 366
 3575 AACGATCAGGTCTGCTGACTCTGCGGAAAAAATCAATG-----CAGATTGTGATAAA 3522
 367 ACCGAGCGGTGATCATCACCCATGGAATGACACCATGGAAGAGACCTTCTTCTCCTC 426
 3521 ACTGACGGTTTTGTTCATCACCCACGTTACAGATACCATGGAAGAACTGCTTATTTCTC 3462
 427 AACCTCAGCGTGAAGAACCAAAACCTGCTGCTTGTAGGCGCCATGGTCCAGGCTCT 486
 3461 GATCTGACCACTCAGTGTCTATAAATCTATCTGTAATGTTGGCCCAATGCTCCATCAAC 3402
 487 TCCATGAGTGTGATGGCCCCCATGAATCTCTATAACCGCGTGAATGTAGCATCAACAA 546
 3401 GCTCTGGTGTGATGGCCCAATGAACCTCTATAATGCTGTTGTTGAGCGTCAGATAAA 3342
 547 GCCTCTACTAAACAAAGGAGTGGTGTGATGATGAGAGTGAAGTCAACGCGCCAGAGAA 606
 3341 GACTCTGCAAAACCGTGGTGTGTTTACTGTGCTATGAATGTTCTGTAATCAAGACGTGAC 3282
 607 GCGACCAAGCTCAACACACACCGCAGTCAATGCTTTCGCCCAACACAGAGTAAATC 666
 3281 ATCGGTAACTCAGCAGCAAACTGAGTTTCAAGTTTCCAGGCTGAGCAGCGGCCACAA 3222
 667 GGCACAGTCTATTATGGCAAGTTCGAGTATTTTCACTCAATCCGTTGACCTCAGCCCTT 726
 3221 GGCCTTCAATCAACCGTAAAGTAACTACTCTCT---GCAGCACCAGTGAAGCAGAT 3165
 727 GCAAGTGAATTTGATATTAAGCAAAATCGAAGAACTCCCGAGTGCATATTTCTTACGCT 786
 3164 AAAGCTGTTTGTGATGTCAGCAAACTGACTGAATCTGCCAAAGTAGGTATTTGTGATAAC 3105
 787 CACCCCGATGATCTGATGTTTGTAGTCAATGCAAGCCCTTCAGGCGAGGCCAAAGGAATC 846
 3104 TACTCTAATGCTTCCGATCTGCGGTTTAAAGCATTTGTAGAGATGATTATTAAGGTATT 3045
 847 ATCCATGCGAGTGGCAATGGAACCTTTCCTTTGACTCAAAATGCTCTTGAAGAA 906
 3044 ATCAGTGCCTGTTGTTGTAACCGTAACTCTATAAGTCTATTCTGGAATCTCTGTCAAA 2985
 907 GCAGCCAAATCAGCGTAGTCTGCTGCTGCAAGCTCTAGAGTGGCGAGTGGTTCACCAACC 966
 2984 GCGCTTAAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925
 967 CAAGCGCTGAAGTGGATGATGAAGAACTTGGTTTGGTTCAGAGAGTCTCAACCTT 1026
 2924 CAAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2865
 1027 CAAAGGCGAGTGGTCTTCTTATGTTAGCCCTCACCAGTCTGATGATGAGAGGCGATC 1086
 2864 CAGAAAGCTGCTGTTCTTCTGCAATGGCTCTGACTCAGACTGCTGACACAGCAAAAT 2805
 1087 CAAAGATCTTCTCCACCTATTATCAAGAA 1118
 2804 CAGAAATGTTTCTAAATATTATCTCTGAA 2773

RESULT 4

US-09-897-516A-1295/c
 ; Sequence 1295, Application US/09897516A

```

; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph B.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1295
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6985)..(9081)
; OTHER INFORMATION:
; US-09-897-516A-1295

```

Query Match	26.7%;	Score 302.8;	DB 5;	Length 9082;
Best Local Similarity	58.1%;	Pred. No. 2.2e-91;		
Matches 576;	Conservative 0;	Mismatches 407;	Indels 9;	Gaps 2;
127	CCCAAAGTGACTATCCTAGGCACAGGAGGCACCATCGCTGTTCCGGGGAACTAGCGTC	186		
3755	CCAAAATTTACTGTTCTGGCAACGGGTGGCACAACTCGCTGGTGGTGAATCTGCAACT	3696		
187	AAGAGTAGTACTCTGCTGGAGCAGTCAACGGTTGATAAGCTTCTTGACGCGCTCCCTGCC	246		
3695	AAATCCAGGTACACTCGGGTCAAGTCGGTGTGATTCAITGCTGAATGCTGTTTCTGCT	3636		
247	ATCAACGACCTAGCCACCATCAAGGTTGACACAGATCTCAAGCAATGGCTCCCAAGAGATG	306		
3635	ATCAAAAACATGCTAATCTGAAGGTGACAGGTTGTTAGCATCGGTTCTCAGGACATG	3576		
307	ACGGGTAAGGTGGCTTAAACTAGCAACGCGTGTCAATGAGCTCTCGGCCCAAAAGAG	366		
3575	AACGATCAGGTCTGGCTGACTCTGGCGAAAAAATCAATG- - - - -CAGATTGTGATAA	3522		
367	ACCGAAGCCGTGATCATACCCATGAACTGACACACATGGAAGAGACCGGTTTCTTCCTC	426		
3521	ACTGACGGTTTGTGCATACCCACCGTACAGATACCATTGGAAGAAAATGCTGTTATTTCTC	3462		
427	AACCTCACGGTGAAGAGCCAAAACCTGCTCTGCCTTGTAGGCGCCCATGGCTCCAGGCTCT	486		
3461	GATCTGACCACCTCAGTGTCTATAAACCTATCGTAATGTTGGCGCAATGGCTCATCAAG	3402		
487	TCCATGAGTCTGATGGCCCATGAATCTCTATAACGCCGTGAATGTAGGATCAACAA	546		
3401	GCTCTGGGTGCTGATGGCCCATGAAACCTCTAATATGCTGTTGTTGTAGCGTCAGATAA	3342		
547	GCCTCTACTAACAAAGGAGTGTGATTGTGATGAACGATGAGATTCAACGCCACAGAA	606		
3341	GACTCTGCAAAACCGTGTGTTTTACTTTGCTATGAATGATTCTGTAATTCATGGACGTGAC	3282		
607	GGACCAAGCTCAACACACCGCAGTCAATGCAATTGGCTTTGGCCCAACACAGGTAAAAATC	666		
3281	ATCGGTAAACTGACACAACTGAAGTTTCAGGCTTTCCAGGCAGTGAACGAGGCCACAA	3222		
667	GGCAGAGTCTATTATGGCAAGTCGAGTATTTTCACTCAATTCGTTGCAGCTCACACCCCT	726		
3221	GGCTTCATTATAACCGGTAAAGTAAACTACTACTCT- - -GCAGCACCTGAAAGCAGAT	3165		
727	GCAAGTGAGTTTGATATTAGCAAAATCGAAGAACTCCCCAGAGTCGATATCTTTACGCT	786		
3164	AAAGCTGTTTGTGTCAGCAAACTGACTGAACCTGCCAAAAGTAGTATTGTGTATAAC	3105		

```

RESULT 5
US-09-897-516A-1296
; Sequence 1296, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1296
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8058)..(9080)
; OTHER INFORMATION:
US-09-897-516A-1296

```

	Query Match	26.7%	Score 302.8	DB 5	Length 9082
	Best Local Similarity	58.1%	Pred. No. 2.2e-91		
	Matches 576	Conservative 0	Mismatches 407	Indels 9	Gaps 2
QY	127	CCCCAAGTGAATCTCTAGCCACAGGAGGCACCATCGCTGGTTCGGGGGAATCTAGCGTC	186		
Db	5328	CCAAACATTACTGTTCTGGCAACGGGTGCACAATCGCTGGTGGTGNATCTGCNACT	5387		
QY	187	AAGAGTAGTACTCTGCTGGAGCAGTACCGTTGATAGCTTCTTCCAGCCGTCCTCGCC	246		
Db	5388	AAATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTCGTAATGCTGTTCCTGCT	5447		
QY	247	ATCAACGACCTAGCCACCATCAAGGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG	306		
Db	5448	ATCAAAACACATTGCTAATCTGAAAGGTGAGCAGGTTGTAGCATCGGTTCTCAGGCATG	5507		
QY	307	ACGGGTAAAGGTGGGCTTTAACTAGCCCAAGCGTGTCAATGAGCTCTCTCCGCCAAAAGAG	366		

5508 AACGATCAGGTCTGGCTGACTCTCGCGGAAAAAATCAATG-----CAGATTGTGATAAA 5561
 QY ACCGAGCCGCTGATCATCACCATGGAATGACACCATGGAAGAGACCGCTTTCTTCCTC 426
 Db ACTGACGGTTTGTTCATCACCCGCTACAGATACCATGGAAGAACTGCTTTCTTCCTC 5621
 QY AACCTCAGGTGAAAGCCAAAAACCTGTCTGCTGTTGAGCGGCAATCGTCCAGGCTCT 486
 Db GATCTGACCACTCAGTGTCTATAAACCTATCGTAATGTTGGCGCAATCGTCCATCAACG 5681
 QY TCCATGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 Db GCTCTGGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5741
 QY GCCTCTACTAACAAAGGAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 606
 Db GATCTGCAACCGTGGTGTGTTTACTTGTCTATGATGATGATGATGATGATGATGATGATGAT 5801
 QY GCGACCAAGCTCAACACACCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 666
 Db ATCGGTAAACTGAGCACAACCTGAAGTTCAGGCTTTCCAGGCGAGTGAACGCGGCGCCACA 5861
 QY GGCACAGTCTATTATGCAAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 726
 Db GGCTTCTATTATGCAAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 5918
 QY GCAAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 Db AAGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5978
 QY CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db TACTCTAATGCTCCGATCTCGCGTTAAAGCAATTTGTAGAGATGATGATGATGATGATGATGAT 6038
 QY ATCATGCGAGCATGGCAATGGCAACCTTCCCTTTGATGATGATGATGATGATGATGATGATGAT 906
 Db ATCAGTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6098
 QY GCAGCCAAATCAGCGGTAGTCTGCTCGCTCGAAGCTCTAGAGTGGCGAGTGGTTCACACCA 966
 Db GCGGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6158
 QY CAAGAGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db CAAAATGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6218
 QY CAAAAGCCAGAGTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
 Db CAGAAAGCTGCTGTTCTTCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6278
 QY CAAAAGATCTTCTCAGCTATTATTAATCAAGAA 1118
 Db CAGAAATGTTTCTAAATATTAATCTGAA 6310

RESULT 6
 US-09-937-516A-1297
 ; Sequence 1297, Application US/09897516A
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Huesing, Joseph E.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spiridonov, Sergei
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)B
 ; CURRENT APPLICATION NUMBER: US/09/897,516A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/215,161
 ; PRIOR FILING DATE: 2000-06-30

; NUMBER OF SEQ ID NOS: 8415
 ; SEQ ID NO 1297
 ; LENGTH: 9082
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (6861)..(7958)
 ; OTHER INFORMATION:
 ; US-09-897-516A-1297
 Query Match 26.7%; Score 302.8; DB 5; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 2.2e-91;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;
 127 CCCCAAGTACTCTCTAGCCACAGGAGGACCATCGCTGTTGCGGGGAATCTAGCGTC 186
 Db CCAACATTTACTTCTGCGAACGGGTGGCACAATCGTGGTGGTGAATCTGCAACT 5387
 QY 187 AAGAGTAGTACTCTGCTGAGCAGTCACCGTTGATAAGCTTCTTGACGCGTCCCTGCC 246
 Db AATCCAGCTACACTGCGGTCAAGTCTGGTTGATTCAATTGCTGAATGCTGTTCTGCT 5447
 QY 247 ATCAACGACTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGCTCCCAAGATG 306
 Db ATCAAAAAACATTTGCTAATCTGAAAGGTGAGCAGGTTGTTAGCATCGGTTCTCAGGACATG 5507
 QY 307 ACGGTGAGGTGCTGCTTAACTAGCCAAAGCGTGTCAATGAGCTCTCTCGCCCAAAAGAG 366
 Db AACGATCAGGTCTGGCTGACTCTGCGGAAAAAATCAATG-----CAGATTGTGATAAA 5561
 QY 367 ACCGAAGCGTGTATCATCAACCATGGAATCTGACACCATGGAAGAGACCGCTTTCTTCCTC 426
 Db ACTGACGGTTTGTCTATCACCCACGCTACAGATACCATGGAAGAACTGCTTTCTTCCTC 5621
 QY 427 AACCTCAGGTGAAAGCCAAAAACCTGCTGCTGCTGTTGAGGCGCATGCGTCCAGGCTCT 486
 Db GATCTGACCACTCAGTGTCTAATGCTGTAATGTTGGCGCAATGCGTCCATCAACG 5681
 QY 487 TCCATGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 Db GCTCTGGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5741
 QY 547 GCCTCTACTAAACAAAGGAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 606
 Db GACTCTGCAACCGTGGTGTGTTTACTTGTCTATGATGATGATGATGATGATGATGATGATGAT 5801
 QY 607 GCGACCAAGCTCAACACACCGCAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 666
 Db ATCGGTAAACTGAGCACAACCTGAAGTTCAGGCTTTCCAGGCGAGTGAACGCGGCGCCACA 5861
 QY 667 GGCACAGTCTATTATGCGCAAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 726
 Db GGCCTTCAATTAACGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5918
 QY 727 GCAAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 Db AAGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5918
 QY 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db TACTCTAATGCTCCGATCTCGCGTTAAAGCAATTTGTAGAGATGATGATGATGATGATGATGATGAT 6038
 QY 847 ATCATGCGAGCATGGCAATGGCAACCTTCCCTTTGATGATGATGATGATGATGATGATGATGAT 906
 Db ATCAGTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6098
 QY GCAGCCAAATCAGCGGTAGTCTGCTCGCTCGAAGCTCTAGAGTGGCGAGTGGTTCACACCA 966
 Db GCGGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6158
 QY CAAGAGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db CAAAATGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6218
 QY CAAAAGCCAGAGTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
 Db CAGAAAGCTGCTGTTCTTCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6278
 QY CAAAAGATCTTCTCAGCTATTATTAATCAAGAA 1118
 Db CAGAAATGTTTCTAAATATTAATCTGAA 6310

Db 6159 CAAAATGCTGAAGTAAATGACAGCAAAATACGGTTTGTGTCATCAGAACGCGCTGAACCCG 6218
QY 1027 CAAAAGCCAGAGTGTCTTATGCTAGCCCTCACCACCAAACTAGTGATAGAGCGGATC 1086
Db 6219 CAGAAAGCTCGTGTCTTCTGCAATGGCTCTGACTCAGACTGCTGACACACACAAAAT 6278
QY 1087 CAAAAGATCTTCCACCTATTAATCCAGAA 1118
Db 6279 CAGGAAATGTTTCTAAATATAATTCCTGAA 6310

RESULT 7
US-09-897-516A-1298
; Sequence 1298, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1298
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5259)..(6302)
; OTHER INFORMATION:
US-09-897-516A-1298

Query Match 26.7%; Score 302.8; DB 5; Length 9082;
Best Local Similarity 58.1%; Pred. No. 2.2e-91;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCAAGTACTATCTTAGCCACACAGGAGGACCATCGCTGTTCGGGGCAATCTAGCGTC 186
Db 5328 CCAAACTACTGTCTGGCAACGGGTGGACATCGCTGGTGGTGAATCTCAACT 5387
QY 187 AAGAGTAGTACTCTCTGTGAGCAGTCCCGTTGATAGCTTCTTGCAGCCGTCCTGCC 246
Db 5388 AATCCAGCTACACTGGGGTCAAGTCGGTGTGATTCTGCTGAATGCTGTTCTGCT 5447
QY 247 ATCAAGACCTAGCCACCATCAAGGTGAACAGATCTCAAGCATGGTCCCAAGAGATG 306
Db 5448 ATCAAAACATCTCTAATCTGAAAGGTGAGCAGGTGTGATGATCGGTCTCAGGACATG 5507
QY 307 ACGGTGAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCTGCCCAAAAGAG 366
Db 5508 AAGCATCAGTCTGGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAA 5561
QY 367 ACGAAGCCGTGATCATACCCATGGAATGACACCATGGAAGAGACCGCTTCTCCTC 426
Db 5562 ACTGACGGTGTGTCATCACCACCGGTACAGATACCATGGAAGAACTGCTTATTTCCTC 5621
QY 427 AACCTCAGGTGAAAGCCAAACCTGCTGCTGCTTGTAGCGCCATCGTCCAGGCTCT 486
Db 5622 GATCTGACCACTCAGTGTGATAAACCTATCGTAATGGTGGCGCAATCGTCCATCAACG 5681
QY 487 TCCATGAGTGTGATGGCCCCCATGAATCTCTATAACGCCGTGAATGAGCGATCAACAAA 546
Db 5682 GCTCGGGTGTGATGGCCCATGAACTCTATATGCTGTGTGTGACGTGAGATATAA 5741
QY 547 GCCTCTACTAAACAAAGAGTGGTGAATGTTGATGAACGATGAGATTCACGCCGCCAGAGAA 606

RESULT 8

US-10-329-670-1

; Sequence 1, Application US/10329670

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, F

; TITLE OF INVENTION: Thereof, and Uses Thereof

; FILE REFERENCE: PB186P1

; CURRENT APPLICATION NUMBER: US/10/329,670

; PRIOR FILING DATE: 2002-12-24

; PRIOR APPLICATION NUMBER: US 09/643,990

; PRIOR FILING DATE: 2000-08-23

; PRIOR APPLICATION NUMBER: US 08/487,429

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/426,787

; PRIOR FILING DATE: 1995-04-21

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 1

; LENGTH: 1830121

; TYPE: DNA

; ORGANISM: Haemophilus influenzae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (4747)..(4747)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (9921)..(9921)

; OTHER INFORMATION: n equals a, t, g or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (10150)..(10150)

```

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10091)..(10091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)
; OTHER INFORMATION: n equals a, t, g or c

Query Match 26.2%; Score 296.4; DB 6; Length 1830121;
Best Local Similarity 57.1%; Pred. No. 4e-88;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
QY 126 ACCCCAGTACATCTCTAGCCACAGGAGGACCATCGCTGGTTCGGGGGATCTAGCGT 185
DB 802722 ACCAAATATTACAATCTTGGCAACGGGTGGTACCATTGCAGGAACGGGCAAGTTCGGT 802781
QY 186 CAAGAGTAGTACTCTCTGAGCAGATCACCCTTGATTAAGCTTCTGCAGCCGCTCCCTGC 245
DB 802782 AATTCTGCTATAAAGCTGGACAAATTAAGTATTGATTAATTAAGCTGTACCAGA 802841
QY 246 CATCAACGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT 305
DB 802842 AATGAAAAATTTGCCAACATTAAGGTGAGCAAAATTTAAAAATAGGTTTCAACAAGACAT 802901
QY 306 GAGGGTAAAGTGTGGCTTAACCTAGCCAGCGTGTCAATGAGCTCCTCGCCCAAAAAGA 365
DB 802902 GAATGACGAAGTCTGGCTAAACCTGGCAAAAGCCATCAATGCTC-----AATGTAAAAG 802955
QY 366 GACCGAAGCGTGTATCATCCCATGGAACCTGGAACCTGGAAGAGACGCTTTCTTCTCT 425
DB 802956 TACTGATGATTTGTCTAATCCCATGTTAGATACCATGGAAGAGAGCGCTTATTCTT 803015
QY 426 CAACCTCACGGTGAAGAACCAAAACCTGTCTGCTTTAGCGGCCATCGCTCCAGGCTC 485
DB 803016 AGATTTAACCGTAAATGTGAAAAACCGTGTGTTCTGTTGGGGCAATCGCTCTGCAAC 803075
QY 486 TTCCATGAGTGTGATGGCCCATGAATCTCTATAACCGCTGAATGTAGCATCAACAA 545
DB 803076 AGAAAAAGTGTGATGGCCCATTAATCTTTTACAATGCTGTCTGTGCGCAGACAA 803135
QY 546 AGCCTCTACTAAACAAAGAGTGGTATTGTGATGAACGATGAGATTTCACGCGCCAGAGA 605
DB 803136 AATATCAATGCTGCTGTTTGTAGTGCATGAATGAATGAATGATGCTGCTGCGCA 803195
QY 606 AGCGCAAGCTCAACACCCAGTCAATGCAATTTGCTTTGCCCAACACAGGTAATAAT 665
DB 803196 TGTAAACAAAACAGTACGACCGCAGTCAAAACGTTCCATTCAACCAATATTGTTCTCT 803255
QY 666 CGGACAGTCTATTATGGCAAGTCGAGTATTCTCAATCCGTTGACCTCACACCT 725
DB 803256 AGGCTATTTCATAACGCAAGTGGACTATGAACGTTTCCCGAGAAAGCAACATACCAT 803315
QY 726 TGAAGTGAAGTTGATATTAGCAAAATCGAAGAACTCCCGCAGAGTCGATATTCTTTACGC 785
DB 803316 CAACATCCGTTTAAACGTAGAAAATTAGATAGCTACCCAAAGTGGGATTTATTATGC 803375
QY 786 TCACCCCGATGATAGTATGTTTATGTAATGTCAGCCCTTCAGGAGGAGGCCAAGGAAT 845

DB 803376 TTATTCAAATGCACCTGTGGAACCATTAACGCAATTACTCAATGTGGTATCAAGGAT 803435
QY 846 CATCCATGAGGATGGCAATGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGAATA 905
DB 803436 TGATCTGAGGAGTGGCAATGGAATGTTAATGCTGCACACTTAGATCGCTTGAATA 803495
QY 906 AGCAGCCAAATCAAGGCGTAGTCTGCTGCTGCAAGCTCTAGAGTGGCAGTGGTTCCACCAC 965
DB 803496 AGCGCAAAAGATAGCGTCTGTTAGTCCGTTCTCTCGTGACCAACGGGTTATACAAC 803555
QY 966 CCAAGAGCTGAAGTGAATGAATAAAGTCTTGGTTTGGTGGCTACAGAGTCTCAACCC 1025
DB 803556 TCGTGACGCTCAAGTGTGATGATAGTAATATAGGCTTTGTAGCATCAGGTACTTTAATCC 803615
QY 1026 TCAAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTAGAGAGAGCGAT 1085
DB 803616 ACAGAAAGCCCGCTGCTCTTGCAATTAGCTTTAACTCAAACTAAAGATCCNAAAGTAAT 803675
QY 1086 CCAAAGATCTTCTCCACCTATTAA 1110
DB 803676 TCAACAATATTTCGAAGACTTCTAA 803700

RESULT 9

US-10-603-114-1219
; Sequence 1219, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1219
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-10-603-114-1219

Query Match 24.5%; Score 277.2; DB 6; Length 1041;
Best Local Similarity 56.5%; Pred. No. 3.7e-83;
Matches 560; Conservative 0; Mismatches 423; Indels 9; Gaps 2;

QY 119 TGGCTAAACCCCAAGTGAATCTCTAGCCACAGGAGGACCAATCGCTGGTTTCGGGGGAAT 178
DB 59 TTGCTTTACCTAATGTGACATCTCGCAACTGGGGGAACCATCGCAGGGGTGGTGACT 118
QY 179 CTAGCGTCAAGAGTAGTACTCTGCTGGAGCAGTCAACGTTGATTAAGCTTTTGCAGCGC 238
DB 119 CTGCCACTTCTTCAAGCTACACCGCAGGAAAATTAGGCATTGATACATTGATTAATGCGC 178
QY 239 TCCTGTCCATCAACGACCTAGCCACCTCAAGGGTGAACAGATCTCAAGCATTTGGTCCC 298
DB 179 TACCTGAGGCAAAAAGTGGCTAATTAACAGGGCAGCAAGTGGTTAATATTGGTTCCG 238
QY 299 AAGAGATGAGGGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATGAGCTCTCGCCC 358
DB 239 AAGATATGAACGACCAAGTGTGGTTAAACTGGCGAATAAATAATTAACG-----GATT 292
QY 359 AAAAAGAGAGCGAAGCGTGTATCAACCCATGGAACCTGACACCATGGAAGACCGCTT 418
DB 293 GTGATAAAACAGACGCGCTTTGTGATACCCATGTTACAGATACGATGGAAGACCGCCT 352
QY 419 TCTTCTCAACCTCACGCTGAAAGCCAAAACCTGTCTGCTCTGTAGCGCCATGGCTC 478
DB 353 ATTTCTTGATTTAAACACAGCATGTAAAAACCGGTGGTGTGATGGGGCAATGGCC 412

QY 479 CAGGCTCTTCCATGAGTGTGATGGCCCCCATGAATCTCTATAACGCCGTGATGAGCGA 538
 Db 413 CAGCAACAGCATTAGGTGCTGATGGTCCATTAAACCTCTATAAATGCGGTGGTGGCAA 472
 QY 539 TCACAAAGCCTCTACTAAACAAGGAGTGGTGTGATGATCAACAGATGAGATTCACGCCG 598
 Db 473 GTGATAAGCATCTGAAATCGTGGTGTATTTAGTTACCATGAATAATTCAGTGTGATG 532
 QY 599 CCAGAGAGCGACCAAGCTCAACACACCGCGAGTCAATGCAATTTGCTTCGCCGCCACACAG 658
 Db 533 GTAAAGATGTTGTCNAATGATACGACAGAGTGCAGCAATTCAGCGGATTAATGAG 592
 QY 659 GTAAATCGCACAGTCTATTATGCGAAAGTCGAGTATTTCACTCAATTCGTTGACCTC 718
 Db 593 GTGCTCAAGGTATGTTTCATGATGGTAAAGTACACTATTACACT---GCTGCAACACCGC 649
 QY 719 ACACCTTGCAAGTGTGATTTGATATTAGCAAAATCGAAGAACTCCCAAGATCGATATTC 778
 Db 650 GCGCTGAGAAAGTTGCTTTGATGTCAGCAAAATTAACCGAATTAACCAAAAGTAGGTATT 709
 QY 779 TTTAGGCTCACCCGATGATCTGATGTTTGTAGTCAATGTCAGCCCTTCAGGCAGGAGCA 838
 Db 710 TTTATACTATGCTAATGATCTGATTTACCGGTAAAGCGTTATAGATTAACCACTTTA 769
 QY 839 AAGGAATPCATCCATGCGAGGATGGCAATGGAAACCCCTTTCCCTTTGATCAAAATGCTC 898
 Db 770 AAGGTATTTAGTGCAGGTGTCGCAATGGTAACTTATATTTGACATCTTAAATACTT 829
 QY 899 TTGAAAAGCAGCAAAATCAGGCTGATGCTGCGCTCGAGCTCTAGAGTGGCAGTGGTT 958
 Db 830 TAGCTGATGTTGTAATAAAGAGTGTGTTGTTGCTTCAAGTCTGTTCCGGTTGGAT 889
 QY 959 CCACCCCAAGAGCGTGAAGTGTGATTAAGAACTTGTGTTGCTGCTACAGAGATC 1018
 Db 890 TTACCAACAAAATGGTGAAGTAGATGATGCAAAATATGGCTTTATTCCTTCAGAACGTT 949
 QY 1019 TCACCTCAAAAGCCAGAGTGTCTTATGTTAGCCCTCACCACAAACTAGTGTATAGAG 1078
 Db 950 TAAATCCACAGAAAGCAAGAGTGTATTACAGTTATCTTTGACAGAAACACAGATCCG 1009
 QY 1079 AGGGATCCAAAGATCTTCCACCTATTAA 1110
 Db 1010 CCAGATCCAAAGAACTTTGAAAAATATTAA 1041

RESULT 10
 US-10-617-320-791
 ; Sequence 791, Application US/10617320
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 ; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS
 ;
 ; NUMBER OF SEQUENCES: 5206
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: <Unknown>
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: <Unknown>
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/617,320
 ; FILING DATE: 10-Jul-2003
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,433
 ; FILING DATE: 30-Jun-1998
 ; APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-011
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 791:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 966 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus pneumoniae
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1...966
 SEQUENCE DESCRIPTION: SEQ ID NO: 791:
 US-10-617-320-791
 Query Match 7.4%; Score 84.4; DB 6; Length 966;
 Best Local Similarity 51.9%; Pred.No. 7e-18;
 Matches 190; Conservative 0; Mismatches 176; Indels 0; Gaps 0;
 QY 369 CCAAGCCGTGATCATCACCCATGGAATGACACCATGGAAGAGACCCGCTTTCTTCTCTCAA 428
 Db 231 CGATGGAGTGGTGTATCACACACGGAACCGATATCTTTAGAGAAACACGCTATTTCTTGA 290
 QY 429 CCTCAGGTGAAAAGCAAAACCTCTGCTGCTTTGTAGGCGCATGCGTCCAGGCTCTTC 488
 Db 291 TACCATGGAAGTTCCCATATGCTATCGTTCTAACAGGAGCCATGCGTAGTCCAATGA 350
 QY 489 CATGAGTGTGATGGCCCATGATCTCTATACGCGGTGAATGTAGCGATCAACAAGC 548
 Db 351 GCTCGTAGTGTGATGTTTATTAATTTACCTAAGTCTTTACGAGTGGCCAGTGTGACAG 410
 QY 549 CTCTACTAAAGAGGAGTGTGATTTGTGATGAACGATGAGATTCACGCCGCCAGAGAGC 608
 Db 411 GGCTGCTGACAAAGAGTGTGCTGCTTATGAACGATGAATCCACGCTGCCAAGTATGT 470
 QY 609 GACCAAGCTCAACACACCGCGAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAATCGG 668
 Db 471 CACCAAAACACATACGACTAATGTCAACACCTTCCAGACTTCCAAACACATGCGCCCTCGG 530
 QY 669 CACAGTCTATTATGGCAAGTGGAGTATTTCACTCAATCGTTCGACCTCACACCTTGC 728
 Db 531 TCTCATGTAACAGGAAATCTCTACTTTCAAAACAGCTGAACCTCGTTCGCTTTGA 590
 QY 729 AAGTGA 734
 Db 591 CCTGA 596
 RESULT 11
 US-10-406-676-3
 ; Sequence 3, Application US/10406676
 ; GENERAL INFORMATION:
 ; APPLICANT: Structural Genomics, Inc.
 ; APPLICANT: Antonyamy, Stephen
 ; APPLICANT: Feil, Ingeborg
 ; APPLICANT: Buchanan, Sean
 ; APPLICANT: Post, Kai W.
 ; APPLICANT: Liu, Yi
 ; APPLICANT: Lorber, David
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF PAK4D KINASE
 ; TITLE OF INVENTION: PAK4KD


```

RESULT 13
PCT-US03-06661A-11
; Sequence 11, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Su-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 10/231,114
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,470
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,063
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,213
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,298
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Streptomyces cinnamonensis ATCC 11874
PCT-US03-06661A-11

```

```

Query Match      5.9%; Score 67.4; DB 1; Length 1074;
Best Local Similarity 98.6%; Pred. No. 4.2e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGGGCGAGCGGCATCATCATCATCATATAGCAGCGGCTGTGTGCCGCGCGCGCAGCCAT 60
      |||
Db      1 ATGGGCGAGCGGCATCATCATCATCATATAGCAGCGGCTGTGTGCCGCGCGCGCAGCCAT 60

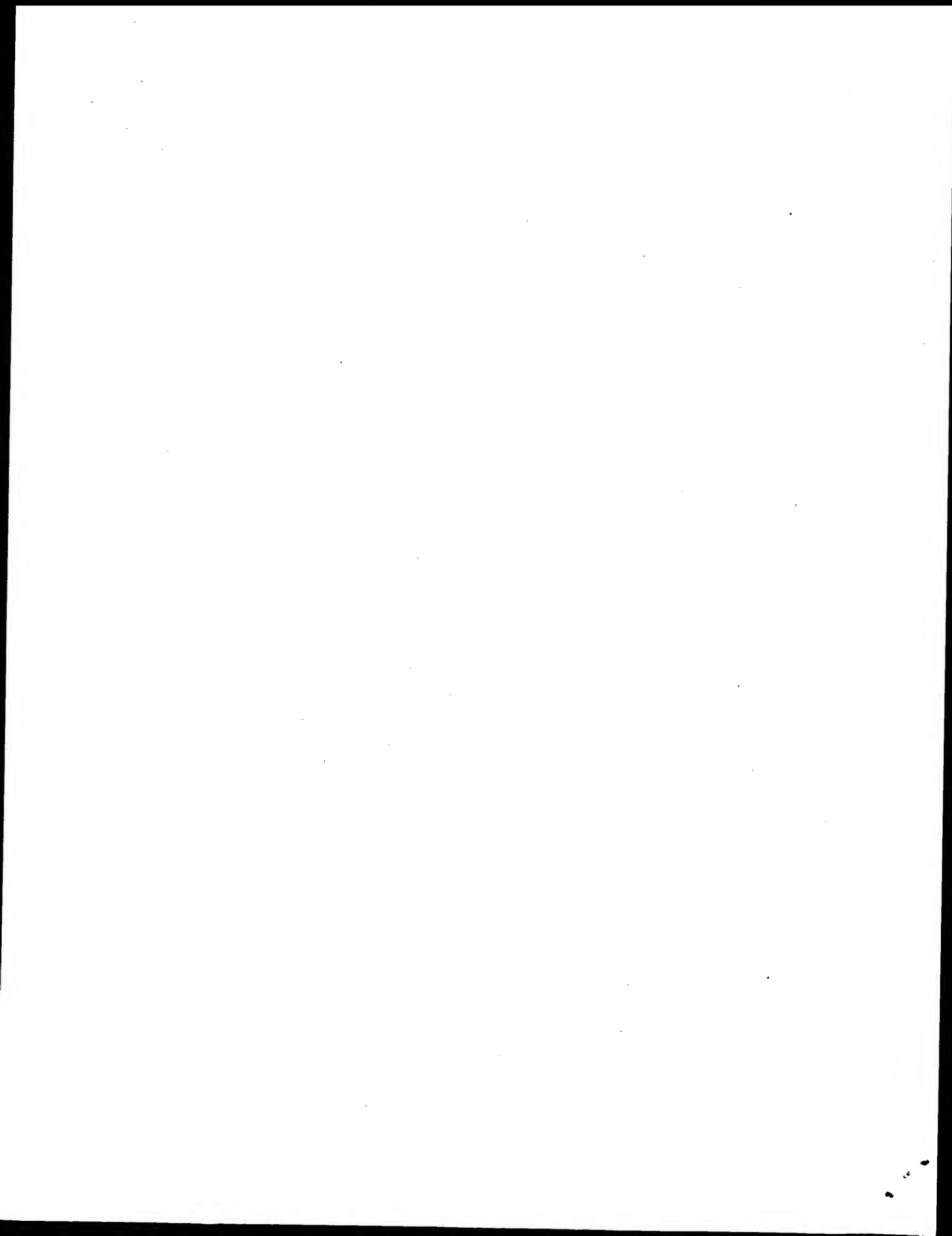
QY      61 ATGGCTAGC 69
      |||
Db      61 ATGGCTAGC 69

RESULT 14
US-10-423-156-60
; Sequence 60, Application US/10423156
; GENERAL INFORMATION:
; APPLICANT: Lin, Hsin-Yu
; APPLICANT: Hwong, Ching-Long
; TITLE OF INVENTION: ANTIGENIC FRAGMENT OF HUMAN
; TITLE OF INVENTION: T-LYMPHOTROPIC VIRUS
; FILE REFERENCE: 05204-020001
; CURRENT APPLICATION NUMBER: US/10/423,156
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: TW 91135980
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated oligonucleotide
US-10-423-156-60

```

Query Match 5.8%; Score 65.2; DB 6; Length 759;
Best Local Similarity 95.7%; Pred. No. 2e-11;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGACGAGCCATCATCATCATCATGACGGCGCTGGTCCGCGCGGACCCAT 60

Search completed: August 27, 2003, 11:03:32
Job time : 341.126 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:33 ; Search time 4824.26 Seconds
(without alignments)
9607.814 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atggcagcagccatcatca.....aagaaagggaatctcttcac 1133

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database :

GenEmbl :

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pin.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1129.8	99.7	1133	6	AR157944	AR157944 Sequence
2	1129.8	99.7	1133	6	AX464427	AX464427 Sequence
3	1027.8	90.7	1035	1	WSANSAGEN	X83689 W.succinoge
4	1027.8	90.7	2505	1	WSDCUANSA	X89215 W.succinoge
5	368.8	32.6	3606	1	WSAJ2933	AP002933 Wolinella
C 6	323.2	28.5	303450	1	AP005085	AP005085 Vibrio pa
C 7	300	26.5	10893	1	AE015308	AE015308 Shigella
C 8	300	26.5	292906	1	AE016988	AE016988 Shigella
C 9	296.8	26.2	10749	1	AE005526	AE005526 Escherich
C 10	296.8	26.2	266658	1	AP002563	AP002563 Escherich
C 11	296.8	26.2	303121	1	AE016766	AE016766 Haemophilus
C 12	296.4	26.2	10173	1	U32758	U32758 Haemophilus
C 13	296.4	26.2	110000	6	AR274513_07	Continuation (8 of
C 14	296.4	26.2	110000	6	AR274513_08	Continuation (9 of
C 15	295.2	26.1	1530	1	ECOLASNI	M34277 E.coli L-as
C 16	295.2	26.1	1643	1	ECOLASBA	M34234 E.coli L-as
C 17	295.2	26.1	10448	1	AE000378	AE000378 Escherich
C 18	295.2	26.1	141744	1	ECU28377	U28377 Escherichia
C 19	294	25.9	1848	6	I73512	I73512 Sequence 15
C 20	294	25.9	1848	12	U06943	U06943 Synthetic c
C 21	282.6	24.9	10769	1	AE013881	AE013881 Versinia
C 22	282.6	24.9	210050	1	AJ414148	AJ414148 Versinia
C 23	276	24.4	22204	1	AE008842	AE008842 Salmonell
C 24	276	24.4	230050	1	AL627277	AL627277 Salmonell
C 25	276	24.4	301574	1	AE016844	AE016844 Salmonell
C 26	273.2	24.1	305961	1	AE016937	AE016937 Bacteroid
C 27	270.4	23.9	10094	1	AE001498	AE001498 Helicobac
C 28	259.2	22.9	1196	6	BD092628	BD092628 Identific
C 29	259.2	22.9	10923	1	AE000585	AE000585 Helicobac
C 30	230.4	20.3	304500	1	AP005953	AP005953 Bradyrhiz
C 31	224.8	19.8	2450	1	ECSEK	X14777 Erwinia chr
C 32	223.2	19.7	2837	6	A14577	A14577 asparaginase
C 33	223.2	19.7	2837	6	E01113	E01113 Genomic DNA
C 34	221	19.5	35000	1	AB000617	AB000617 Bacillus
C 35	221	19.5	221160	1	BSUB0002	Z99105 Bacillus su
C 36	215.2	19.0	1399	1	ECASN	X12746 Erwinia chr
C 37	210.8	18.6	1017	6	AR177619	AR177619 Sequence
C 38	209.2	18.5	1709	1	AF056495	AF056495 Pseudomon
C 39	207.6	18.3	314150	1	CJ11168X1	AL139074 Campyloba
C 40	205.2	18.1	308015	1	AE016783	AE016783 Pseudomon
C 41	203.2	17.9	110000	2	AC020884_1	Continuation (2 of
C 42	200.4	17.7	9937	1	AE004563	AE004563 Pseudomon
C 43	195.4	17.2	2353	1	PFL238710	AJ238710 Pseudomon
C 44	178.8	15.8	20516	1	AE008866	AE008866 Salmonell
C 45	128.8	11.4	1337	8	SPLASPA	Y11944 S.pombe gen

ALIGNMENTS

RESULT 1

AR157944

LOCUS

DEFINITION

AR157944

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AR157944

Sequence

3 from patent US 6251388.

AR157944

AR157944.1

GI:16219888

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1133)

Durden,D.L.

Utilization of Wolinella succinogenes asparaginase to treat

diseases associated with asparagine dependence

Patent: US 6251388-A 3 26-JUN-2001;

linear

DNA

1133 bp

PAT 17-OCT-2001

421	TTCTCAACCTCAGCGTGAAAGAGCAAAACCTGTGCTGTAGGCGCCATGCGTCCA	480	/clone_lib="lambda EMBL3"			
481	GGCTCTTCATGAGTGTGATGCGCCCATGAATCTCTATAACGCGGTGAATGATGCGATC	540	20..1012			
481	GGCTCTTCATGAGTGTGATGCGCCCATGAATCTCTATAACGCGGTGAATGATGCGATC	540	/gene="ansA"			
541	AACAAAGCCTCTACTAACAAAGAGTGGTGATGTGATGAACGATGAGATTACGCGCGC	600	20..1012			
541	AACAAAGCCTCTACTAACAAAGAGTGGTGATGTGATGAACGATGAGATTACGCGCGC	600	/gene="ansA"			
601	AGAGAAGCGCAAGCTCAACACACGCGAGTCAATGCTTCGCGCCCAACACAGGT	660	/EC number="3.5.1.1"			
601	AGAGAAGCGCAAGCTCAACACACGCGAGTCAATGCTTCGCGCCCAACACAGGT	660	/codon_start=1			
661	AAAAATCGGCACAGTCTATTATGGCAAGTGGTCAATGCTTCGCGCCCAACACAGGT	720	/transl_table=11			
661	AAAAATCGGCACAGTCTATTATGGCAAGTGGTCAATGCTTCGCGCCCAACACAGGT	720	/product="L-asparaginase"			
721	ACCTTCGCAAGTGGTGGTATATAGCAAAATCGAAGTCTCCCGAGTTCGATATCTT	780	/protein_id="CAA58658.1"			
721	ACCTTCGCAAGTGGTGGTATATAGCAAAATCGAAGTCTCCCGAGTTCGATATCTT	780	/db_xref="GI:758652"			
781	TACGCTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840	/translation="MAKPOVITLAIAGSBSSVKSSYSAGAVTVDKLLAAVPAI			
781	TACGCTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840	NDLTIKGSQISSIGSEMTGKWLKASVNNELAKVETAVIITHTGDTMEETAFPI			
841	GGAATCATCATGACGCGATGGCAATGGGAACCCCTTCCCTTTGACTCAAAATGCTCTT	900	LNLTVKSKPVLVGMARPGSSMSADGPNLNVNVAINKASTNKGVIIVMNDIHA			
841	GGAATCATCATGACGCGATGGCAATGGGAACCCCTTCCCTTTGACTCAAAATGCTCTT	900	AREATKLNNTAVNAPASGNTKIGTYIKGVYFTQSVRPHLASEPDIKIELPRV			
901	GAAGAAGCAGCAAAATCAGGCGTGTCTGCTCGAAGTCTAGAGTGGGCGAGTGGTTC	960	DILYAPDDTDLVNAALOGAKGIHAGMGNGPNPFLQNALAKSAGVGVVAVRSR			
901	GAAGAAGCAGCAAAATCAGGCGTGTCTGCTCGAAGTCTAGAGTGGGCGAGTGGTTC	960	VSGSTTQBEAEVDKLGTFVATESLNPKARVLLMLTKTSDREALQKIFSTY"			
961	ACCACCCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020				
961	ACCACCCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020				
1021	AACCTCAAAAGCAGAGTGTCTTATGTTAGCCCTCACAAAATAGTGTATGAGAG	1080				
1021	AACCTCAAAAGCAGAGTGTCTTATGTTAGCCCTCACAAAATAGTGTATGAGAG	1080				
1081	GGATCCAAAGATCTCTCCACCTATTATTAATCCAAAGAGGGAATCTCTTCCAC	1133				
1081	GGATCCAAAGATCTCTCTCCACCTATTATTAATCCAAAGAGGGAATCTCTTCCAC	1133				
RESULT 3						
WSANSAGEN						
LOCUS	WSANSAGEN	1035 bp	DNA linear BCT 01-APR-1995			
DEFINITION	W.succinogenes ansA gene.					
ACCESSION	X83689					
VERSION	X83689.1	GI:758651				
KEYWORDS	ansA gene; asparaginase.					
SOURCE	Wolinella succinogenes					
ORGANISM	Wolinella succinogenes					
	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;					
	Helicobacteraceae; Wolinella.					
REFERENCE	1					
AUTHORS	Derst, C. and Roehm, K.H.					
TITLE	Cloning and sequencing of L-asparaginase from Wolinella					
JOURNAL	succinogenes					
REFERENCE	2 (bases 1 to 1035)					
AUTHORS	Roehm, K.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (05-JAN-1995) K.H. Roehm, Physiologische Chemie der					
FEATURES	Philippe-Univ.; Karl-von-Friesch-Str.1, D-35033 Marburg (Lahn), FRG					
	Location/Qualifiers					
	1..1035					
	/organism="Wolinella succinogenes"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:844"					
BASE COUNT 293 a 277 c 236 g 229 t						
ORIGIN						
Query Match 90.7%; Score 1027.8; DB 1; Length 1035;						
Best Local Similarity 99.8%; Pred. No. 1e-284;						
Matches 1029; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
QY	103	AGCGCTCTGTTTGTGATGGCTAAACCCCAAGTACTATCTCTGCTGGAGCATCACCGTTCAT	162			
DB	5	AGCGCTCTGTTTGTGATGGCTAAACCCCAAGTACTATCTCTGCTGGAGCATCACCGTTCAT	64			
QY	163	GCTGTTGGGGGAATCTAGCGTCAAGAGTAGTACTCTCTGCTGGAGCATCACCGTTCAT	222			
DB	65	GCTGTTGGGGGAATCTAGCGTCAAGAGTAGTACTCTCTGCTGGAGCATCACCGTTCAT	124			
QY	223	AAGCTTCTTTCGAGCGTCCCTCCCATCAACGACCTAGCACCATCAAGGTTCAACAGATC	282			
DB	125	AAGCTTCTTTCGAGCGTCCCTCCCATCAACGACCTAGCACCATCAAGGTTCAACAGATC	184			
QY	283	TCAGCATTTGGTCTCCCAAGAGTAGCGGTAAAGTGTGGCTTAAAGTAAAGTAAAGTAAAG	342			
DB	185	TCAGCATTTGGTCTCCCAAGAGTAGCGGTAAAGTGTGGCTTAAAGTAAAGTAAAGTAAAG	244			
QY	343	AATGAGTCTCTCGCCCAAAAGAGACCGAAGCGGTGATCATCAACCATGGAATGACAC	402			
DB	245	AATGAGTCTCTCGCCCAAAAGAGACCGAAGCGGTGATCATCAACCATGGAATGACAC	304			
QY	403	ATGGAAGAGACCGCTTCTCTCAACCTCAGCGTCAAGGCAAGGCAAGGCAAGGCAAGG	462			
DB	305	ATGGAAGAGACCGCTTCTCTCAACCTCAGCGTCAAGGCAAGGCAAGGCAAGGCAAGG	364			
QY	463	GTAGGCGCATGCGTCCAGGCTCTTCCATGAGTGTGATGGCGCCCATGATCTCTATAAC	522			
DB	365	GTAGGCGCATGCGTCCAGGCTCTTCCATGAGTGTGATGGCGCCCATGATCTCTATAAC	424			
QY	523	GCGTGAATGTAGCGATCAACAAAGCTCTACTAAACAAAGAGTGGTGGTGGTGGTGGT	582			
DB	425	GCGTGAATGTAGCGATCAACAAAGCTCTACTAAACAAAGAGTGGTGGTGGTGGTGGT	484			
QY	583	GATGAGATTCAAGCGCGCAGAGAGCGCAAGCTCAACACCGCAGTCAATGCAATTT	642			
DB	485	GATGAGATTCAAGCGCGCAGAGAGCGCAAGCTCAACACCGCAGTCAATGCAATTT	544			
QY	643	GCTTCGCCCAACACAGGTAAATCGGCACAGTCTATTATGGCAAGTGGATTTTCACT	702			
DB	545	GCTTCGCCCAACACAGGTAAATCGGCACAGTCTATTATGGCAAGTGGATTTTCACT	604			
QY	703	CAATCGGTTCGACTCACACCTTCAGTCTTGTGATATTATAGCAAAATCGAAGAACTC	762			
DB	605	CAATCGGTTCGACTCACACCTTCAGTCTTGTGATATTATAGCAAAATCGAAGAACTC	664			
QY	763	CCAGAGTGGATATTCTTTAGCTCACCCCGATGATCTGATGTTTGTAGTCAATGCAGCC	822			
DB	665	CCAGAGTGGATATTCTTTAGCTCACCCCGATGATCTGATGTTTGTAGTCAATGCAGCC	724			

QY	703	CAATCGGTTGCACTCACACCTTGCAGTGAAGTTTGATATTAGCAAAATCGAAGAACTC	762	
Db	1774	CAATCGGTTGCACTCACACCTTGCAGTGAAGTTTGATATTAGCAAAATCGAAGAACTC	1833	
QY	763	CCAGAGTCGATATCTTTTACGCTCACCCCGATGATAGTATGTTTATAGTCAATCGAGCC	822	
Db	1834	CCAGAGTCGATATCTTTTACGCTCACCCCGATGATAGTATGTTTATAGTCAATCGAGCC	1893	
QY	823	CTTCAGGAGGAGCCAAAGGAATCATCTCATGAGCATGGCAATGGGAACCTTTCCCT	882	
Db	1894	CTTCAGGAGGAGCCAAAGGAATCATCTCATGAGCATGGCAATGGGAACCTTTCCCT	1953	
QY	883	TTGACTCAAAATGCTCTTTGAAAAAGCAGCCAAATCAGGCGTAGTCTGCTCGAAGCTCT	942	
Db	1954	TTGACTCAAAATGCTCTTTGAAAAAGCAGCCAAATCAGGCGTAGTCTGCTCGAAGCTCT	2013	
QY	943	AGAGTGGGAGTGGTTTCCACCACCCAGAGGCTGAAGTGGATGATAGAACTTTGGTTTT	1002	
Db	2014	AGAGTGGGAGTGGTTTCCACCACCCAGAGGCTGAAGTGGATGATAGAACTTTGGTTTT	2073	
QY	1003	GTGGCTACAGAGTCTCAACCTTCAAAAGCCAGAGTGTCTTCTATGTTAGCCCTCAC	1062	
Db	2074	GTGGCTACAGAGTCTCAACCTTCAAAAGCCAGAGTGTCTTCTATGTTAGCCCTCAC	2133	
QY	1063	AAAACTAGTGATAGAGCGGATCCAAAAGATCTTCTCCACTATTAAATCCAGAAAGG	1122	
Db	2134	AAAACTAGTGATAGAGCGGATCCAAAAGATCTTCTCCACTATTAAATCCAGAAAGG	2193	
QY	1123	AATCTCTTAC	1133	
Db	2194	AATCTCTTAC	2204	
RESULT 5	WSAJ2933	3606 bp DNA linear	BCT 05-OCT-2000	
LOCUS	Wolinella succinogenes aspA, dcuA genes and partial ansA gene.			
DEFINITION	Wolinella succinogenes aspA, dcuA genes and partial ansA gene.			
ACCESSION	AJ002933			
VERSION	AJ002933.1	GI:2644958		
KEYWORDS	ansA gene; aspA gene; aspartate ammonia-lyase; C4-dicarboxylate membrane transporter; dcuA gene; L-asparaginase.			
SOURCE	Wolinella succinogenes			
ORGANISM	Wolinella succinogenes			
REFERENCE	1			
AUTHORS	Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.			
TITLE	Transport of C(4)-dicarboxylates in Wolinella succinogenes			
JOURNAL	J. Bacteriol. 182 (20), 5757-5764 (2000)			
MEDLINE	20461222			
PUBMED	11004174			
REFERENCE	2			
AUTHORS	Ullmann, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (19-NOV-1997) Ullmann R., Institut fuer Mikrobiologie, Johann-Wolfgang-Goethe Universitaet, Biozentrum N240, Marie-Curie-Strasse 9, 60439 Frankfurt, GERMANY			
COMMENT	Related sequence: X89215.			
FEATURES	Location/Qualifiers			
source	1..3606			
	/organism="Wolinella succinogenes"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:844"			
gene	400..1815			
	/gene="aspA"			
CDS	400..1815			
	/gene="aspA"			
	/codon_start=1			
	/transl_table=11			
	/product="aspartate ammonia-lyase"			
	/protein_id="CAA05764.1"			
	/db_xref="GI:2644959"			
db_xref="SPTREMBL:O34244"				
/translation="MASTRIEHLIDGLEVSNDCYVGVQVOTAKENFHISGVLSSMP				
TFIAGLAKVAAALANFELGLLEKIKNAICEACDLIIAGKYHQFVVDALIQGAGT				
STNNANEVIANVALELGHKKVEYKCHPNHNVLSOSTNDAYTAFVVALYEKLIE				
LTVSMGILRDSAPKAEFKDIIKMGRTQLODAPVMTLQGEFTKVTAMVEDIDRVLE				
ANRLVREMLGATAIGTGINSHPDIATLVEKLOQVTPGPFITASDLVEATQDTGAYV				
QISGVLRVSTKYSKICNDRLSSGRAGLSINLIPNQPGSSIPMGKVNPIPRV				
NOVCVHIGADMTVTIASEGQOLNPFVPIAFSSINMMRRAFETLAEKCVNGI				
TANADNCRDYLKSLIGLVTALNPYIGYENSTVAKALESGSKSVHDI VLERKLLSKEE				
LEELIRPENMIKPKVVLTKK"				
1828..3129				
/gene="dcuA"				
1828..3129				
/gene="dcuA"				
/codon_start=1				
/transl_table=11				
/product="C4-dicarboxylate membrane transporter"				
/protein_id="CAA05765.1"				
/db_xref="GI:2644960"				
/db_xref="SWISS-PROT:O34245"				
/translation="MLVFEILVLAALFLGVRLLGIIAGVAGGLMIVLCGLGLKPG				
SIPDIVLIIIMSVIAAIANMQVAGGLDVLVHLAEKLRSPKHITFLAPVVTVMETLL				
AGTGHTAFSTLPVIAKVEQKVPSPSLAVVASQIALITASPVSAAVFMGALBP				
LGVYLOLLAICIPTTFIGCMITAFICNLFNLDKDPVYOERVAKGVMKLRGOTQV				
AKPKAKLSVLIFLVGLAVFYATAISKNVGLIQNPVGRDSAIMMFLMTATLAF				
AKIDIDAVLNASTFKSGWTACICVLGAWLGDTEVSNHNDIKGFAGGILFEGHPMLA				
ITLFFASMLLYSQATKALIPAAIALSVDPVTLLIASFAAVSALFVLPTPTLLAAVQ				
MDDTGSTRIGKVFVNHPIVPGVLAIGISVALGFIVAPILL"				
3250..3606				
/gene="ansA"				
3250..3606				
/gene="ansA"				
/codon_start=1				
/transl_table=11				
/product="L-asparaginase"				
/protein_id="CAA05766.1"				
/db_xref="GI:2644961"				
/translation="MAKPOVTILATGTTAGTGGSSVSKSSVSAGAVTVDKLLAAVPAI				
NDLATIKGEQISSIGSGEOMTGKWLKAKRVNLLAQKETEAVIIITHGTDTMEETAFF				
LNLTVKSRQKPVVLVGM"				
BASE COUNT	919 a	973 c	806 g	908 t
ORIGIN				
Query Match	32.6%;	Score 368.8;	DB 1;	Length 3606;
Best Local Similarity	99.5%;	Pred. No. 1.2e-94;		
Matches	370;	Conservative	0;	Mismatches 2;
Indels	0;	Gaps	0;	
QY	103	AGCGCTCTGTTTGTATGGCTAAACCCCAAGTCACTATCTAGCCACAGGAGGACCATC	162	
Db	3235	AGCGCTCTGTTTGTATGGCTAAACCCCAAGTCACTATCTAGCCACAGGAGGACCATC	3294	
QY	163	GCTGTTTCGGGGGAATCTAGCGTCAAGAGTAGCTACTCTCTGGAGAGTCACTGTTGAT	222	
Db	3295	GCTGTTTCGGGGGAATCTAGCGTCAAGAGTAGCTACTCTCTGGAGAGTCACTGTTGAT	3354	
QY	223	AAGCTTCTTCAGCGCTCCCTGCCATCAACGACCTAGCCACCATCAAGGGTGAACAGATC	282	
Db	3355	AAGCTTCTTCAGCGCTCCCTGCCATCAACGACCTAGCCACCATCAAGGGTGAACAGATC	3414	
QY	283	TCAGCATTTGCTCCCAAGAGATGACGGGTAAAGTGTGGCTTAAATAGCCAGCGTGTCT	342	
Db	3415	TCAGCATTTGCTCCCAAGAGATGACGGGTAAAGTGTGGCTTAAATAGCCAGCGTGTCT	3474	
QY	343	AATGAGTCTCTCGCCCAAAAAGAGACCGAAGCGGTGATCATCACCATGGAACCTGACACC	402	
Db	3475	AATGAGTCTCTCGCCCAAAAAGAGACCGAAGCGGTGATCATCACCATGGAACCTGACACC	3534	
QY	403	ATGGAAGAGACCGCTTCTTCTCAACCTCACGCTGAAAAGCAAAAACTGTCTGCTT	462	
Db	3535	ATGGAAGAGACCGCTTCTTCTCAACCTCACGCTGAAAAGCAAAAACTGTCTGCTT	3594	
QY	463	GTAGGCGCCATG	474	

Db	3595	GTAGGCCCATG	3606	gene	complement (1835. .2314)
RESULT 6				CDS	/gene="VPA0301"
LOCUS	AP005085	303450 bp	DNA		complement (1835. .2314)
DEFINITION	Vibrio parahaemolyticus DNA, linear	BCT 05-MAR-2003			/gene="VPA0301"
ACCESSION	AP005085	BA000032			/note="similar to DBJ:BA043758.1 (AP003138) percent identity 57 in 158 aa"
VERSION	AP005085.1	GI:28808465			/codon_start=1
KEYWORDS					/evidence=not experimental
SOURCE	Vibrio parahaemolyticus				/transl_table=11
ORGANISM	Vibrio parahaemolyticus				/product="putative peptide methionine sulfoxide reductase"
REFERENCE	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.				/protein_id="BAC61644.1"
AUTHORS	Nasu, H., Iida, T., Sugahara, T., Yamauchi, Y., Park, K.S., Yokoyama, K., Makino, K., Shinagawa, H. and Honda, T.				/db_xref="GI:28808467"
TITLE	A filamentous phage associated with recent pandemic Vibrio parahaemolyticus O3:K6 strains				/translation="MMKIKYFAGCGLMGVQFMRLHFGVISTEAGRANGKTDNTQSEY DGVAECVTRTFEDPSQVSIETLMGFFIIPYSINKQDDVYGERKTVGVRDTHKLE IAKSFIAIRADADKIAVELPLKNYVPSDEEHQDLTRFPNDYCHILPDLHLKYNK" 2493. .2882
JOURNAL	J. Clin. Microbiol. 38 (6), 2156-2161 (2000)				/gene="VPA0302"
MEDLINE	20295086				/note="similar to GB:AAA64344.1 (L16865) percent identity 52 in 129 aa"
PUBMED	10834969				/gene="VPA0302"
REFERENCE					/codon_start=1
AUTHORS	Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A., Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M. and Iida, T.				/evidence=not experimental
TITLE	Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae				/transl_table=11
JOURNAL	Lancet 361 (9359), 743-749 (2003)				/product="hypothetical protein"
MEDLINE	22508454				/protein_id="BAC61645.1"
PUBMED	12620739				/db_xref="GI:28808468"
REFERENCE					/translation="MKIEHVAIWTRELELKGFEYKFNVAVSNKYHNPKKHFFSSYFL SPESGARLEMSMEGVTTGNSHMQVTGLAHFAFALGSEQAVDQITKLIVEDGYQRI DGPRTYGDGYESCVDLPDGNRIELTV" 2974. .3714
AUTHORS	Oshima, K., Kurokawa, K., Makino, K., Yokoyama, K., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M. and Iida, T.				/gene="VPA0303"
TITLE	Direct Submission				/note="similar to GB:AAL20622.1 (AE008775) percent identity 54 in 243 aa"
JOURNAL	Submitted (09-APR-2002) Ken Kurokawa, Osaka University, Genome Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871, URL: http://www.gen-info.osaka-u.ac.jp/, Fax: 81-6-6879-2047				/codon_start=1
COMMENT	Genome project				/evidence=not experimental
FEATURES	This clone was isolated from a patient presenting with acute gastroenteritis.				/transl_table=11
source	Location/Qualifiers				/product="putative regulatory protein"
	1. 303450				/protein_id="BAC61646.1"
	/organism="Vibrio parahaemolyticus"				/db_xref="GI:28808469"
	/mol_type="genomic DNA"				/translation="MNPONEILQVNDKRVQVTELSIIIGSVGTIRQDLNLFLEQQ GYLKRVHCAATALQSDDIITRELFKFDIKLANKADLVAPNETVLIEGGSANALLA RTLAERGDVITITSAIHLIRNTSANIILGGVYHQGESLVGLTKLCIENTHFS TAFGLDGFQDQFTSRDMRADIAEAILAKKRNIVLTDSSKFGQIYPPSSIGKTNE ISVLLTDAAPKSLDLEQLKLGVEVILG" complement (3733. .4089)
	/sub_strain="RIMD 2210633"				/gene="VPA0304"
	/db_xref="taxon:670"				complement (3733. .4089)
	81. .1574				/gene="VPA0304"
	/chromosome="2"				/note="similar to GB:AAF96321.1 (AE004376) percent identity 76 in 117 aa"
gene	81. .1574				/codon_start=1
CDS	/gene="VPA0300"				/evidence=not experimental
	/note="similar to EMB:CAC46187.1 (AL591787) percent identity 68 in 487 aa"				/transl_table=11
	/codon_start=1				/product="conserved hypothetical protein"
	/transl_table=11				/protein_id="BAC61647.1"
	/product="hypothetical signal peptide protein"				/db_xref="GI:28808470"
	/protein_id="BAC61643.1"				/translation="MHQHEKLVNVEFTGNIGATKAFEFHVFQWQFVDYGPDYAASFSG QSIDGFIYAEQVQTSNGALLVFYSSDIHATLEKXVAKFGQIIRPIPEFGGCRPH FTPEPSGNEFAVWSESH" 4319. .5860
	/db_xref="GI:28808466"				/gene="VPA0305"
	/translations="MKIKAVPKALLVTLASSLPISVAYAPNANESSILVQOQWEGN WPTQNAEKLVQELYYQRAISAYQPMNVIGMRDGSFKGKGVNVLPIWKQMD RSLVPTPNADVVYNNYLDLKEITGLVVAAPPKVIQMTDFQFTITDVGILIGDRA RGLYLPLPNVDGVPKGYFTFTSTNNVLFERTVMKKGNGADPSDAVKIAQRT IYPLWDEENIQPMKFPNAGORINMVPDTSYWEKMQFVDEEPIAISPELRGVL ASGIVKGFNPTVEKOLLEKAVKTAQKTLGLKRDKNRLYYKQRYERAWA GATSEYMSYLLIDQRAAPFYQYSSAPAMVMTNAGSKYPTVRDSKGDILNGSH QYKHLHPANPAKLPWAVTLNVVDGTPETSQLPSKNGFDNIENKTDGSDILYFGP QLPNGAPESNYIKTIPGRDFLTAIRLYGTGIEFFQTKWPKDDVLKIK" gene				/note="similar to REF:NP_231225.1 (NC_002505) percent identity 83 in 510 aa"
	/codon_start=1				/evidence=not experimental
	/transl_table=11				/product="catalase"
	/protein_id="BAC61648.1"				

71022	Db	ATGACCAACATTGCGGATATTTCTGGTGAGCAAGTCTGACGATCCGGCTCTCTCAAGACATG	70936
307	Qy	ACGGGTAAAGTGTGGCTTTAAACTAGCCCAAGCGTCTCAATGAGCTCTCGCCCAAAAAAGAG	366
70962	Db	AACGATGAGGTATGGCTGAAGTTAGCGAAAGGGTTAACGAGCTGCTCGCTCAAGATGAT	70903
367	Qy	ACCGAAGCCGTGATCATCACCCATCGGAACCTGACACCATGGAAGAGACGGCTTTCTTCCTC	426
70902	Db	GTTGACGGTATCGTTCATCACCCACGCTACGGATACGTTTGGGAAGAACCGCATACTTCTT	70843
427	Qy	AACCTCAGGTGAAAGCCAAACAACTCTCTGCCCTGTAGCGCGCATCGCTCCAGGCTCT	486
70842	Db	GATTTAAACGGTGAAGCGGACAAGCCGGTTGTGATCGTGGCGCGATGGACCATCCACC	70783
487	Qy	TCCATGAGTGTGATGGGCCCAATGAATCTCTATAAAGCGCGTGAATGTAGCGATCAACAAA	546
70782	Db	GCCATGAGTGGCGATGGCCCGGTCAATTTATCAATGCTGTGTTACTGCAACCCGATGAA	70723
547	Qy	GCCTCTACTTAAACAAGGAGTGTGATGTGAACGATGAGATTCAGCGCCGCCAGAGAA	606
70722	Db	GACTCAAAAGCGCGCGCGTGTAGTGGAATGAATGACACTATTTTCGATCGCGGAGAT	70666
607	Qy	GCACCAAGCTCAACACCAACCGGAGTCAATGCAATTTGCTTGGCCCAACACAGTAAAAATC	666
70662	Db	GTAACGAAGACCAACACCACTCAGTAAGTAGCTTCCAATCGCAAACTTCGGCCCTCTT	70603
667	Qy	GGCACAGTCTATTATGGCAAGTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCTTT	726
70602	Db	GGCTACATTCACAACAGCGATGCAAGTAGTACCAAGAAGCCCTGAACGTAAACACACACA	70544
727	Qy	GCAAGTGAGTGTGATATTAGCAAAATTCGAAGAACTCCACAGTTCGATATCTTTTACGCT	786
70542	Db	GAAACCCCAATTCGACGTATCAAAACTCAACACTACCTTAAAGTAGCGATTTGTTTAAAC	70488
787	Qy	CACCCCGATGATCTGATGTTTTAGTCAATGACAGCCCTTCAGGCAGGAGCCAAAGGAATC	846
70482	Db	TATCGGAATCGGTTCGAGCTTCCGAGTCAAGACCTTGGTTGATGAGCTGCGGATTT	7042
847	Qy	ATCCATGACGAGCATGGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA	906
70422	Db	GTGAGTGTGGTGTGGGTAAATGTTGTAACACACGGTCTTTGATGAGCTCGAAAAA	70366
907	Qy	GCAGCCAANTCAGCGTAGTCGTCCGCTCGAAGCTCTAGAGTGGCGAGTGGTTCCACCAAC	966
70362	Db	CGAGCAAGATGGCAATTTGTTGTCGACGAAGCTCTCGACGCGCAACTGGCTCAACGACT	70303
967	Qy	CAAGAGGCTGAAGTGGATGATAAGAACTTGTTGTTGTGGCTACAGAGAGTCTCAACCGT	10266
70302	Db	TTGGACGCGGAAATTGATGATGATATAACGGCTTTGTTGTCATCAGGACGCTAAACCCA	7024
1027	Qy	CAAAAGCCAGAGTCTTCTTTATGTTAGCCCTTCACAAACACTAGTGTATGAGAGCGCATC	1088
70242	Db	CAAAAGCAACGATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	7018
1087	Qy	CAAAAGATCTTCTCCACCTATTAA	1110
70182	Db	CAGAAGATCTCCAGTACTACTTAA	70159

RESULT 7	10893 bp	DNA	linear	BCT 18-OCT-2001
AE015308	10893 bp	DNA	linear	BCT 18-OCT-2001
AE015308/c				
LOCUS	Shigella flexneri 2a str. 301			
DEFINITION	Shigella flexneri 2a str. 301 of the complete genome.			
ACCESSION	AE015308			
VERSION	AE015308.1			
KEYWORDS	GI:24053358			
SOURCE	Shigella flexneri 2a str. 301			
ORGANISM	Shigella flexneri 2a str. 301			
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Shigella.			
REFERENCE	1 (bases 1 to 10893)			

	Query Match	28.58;	Score 323.2;	DB 1;	Length 303450;
	Best local Similarity	58.08;	Pred. No. 3.4e-81;		
	Matches 571;	Conservative 0;	Mismatches 413;	Indels 0;	Gaps 0;
127	CCCAAGTGACTATCTTAGCCACAGGAGGACCAATCCTGTTCCGGGGGAATCTAGCGTC	186			
71142	CCTAATATCAAAATCCTCGACTGGCGGACCATAGCTGGCGGAGGTCAATCTGCCACT	71083			
187	AAGAGTAGCTACTCTGTGGAGAGTACCCTTGATAGCTTCTTGACGGCGTCCCTGCC	246			
71082	GAGTCGAAGTACACCGCAGGTAAAGTTGGCGTCGAATCATTAATTTCTGTGTGCCAGC	71023			
247	ATCAACGACCTTAGCCACCATCAAGGTTGAAACAGATCTCAAGCATTTGGCTCCCAAGAGATG	306			

AUTHORS

Jin, Q., Yuan, Z.H., Xu, J.G., Wang, Y., Shen, Y., Lu, W.C., Wang, J.H., Liu, H., Yang, J., Yang, F., Ou, D., Zhang, X.B., Zhang, J.Y., Yang, G.W., Wu, H.T., Dong, J., Sun, L.L., Xue, Y., Zhao, A.L., Gao, Y.S., Zhu, J.P., Kan, B., Chen, S.X., Yao, Z.J., He, B.K., Chen, R.S., Ma, D.L., Qiang, B.Q., Wen, Y.M., Hou, Y.D. and Yu, J.
Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157
Nucleic Acids Res. 30 (20), 4432-4441 (2002)
2 (bases 1 to 10893)
Jin, Q., Shen, Y., Wang, J.H., Liu, H., Yang, J., Yang, F., Zhang, X.B., Zhang, J.Y., Yang, G.W., Wu, H.T., Dong, J., Sun, L.L., Xue, Y., Zhao, A.L., Gao, Y.S., Zhu, J.P., Chen, S.X., Yao, Z.J., Wang, Y., Lu, W.C., Qiang, B.Q., Wen, Y.M. and Hou, Y.D.
Direct Submission
Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry of Public Health, 100 YingXin Jie, XuanWu Qu, Beijing 100052, P.R. China

TITLE

Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1. .10893

/organism="Shigella flexneri 2a str. 301"

/mol_type="genomic DNA"

/strain="301"

/serotype="2a"

/db_xref="taxon:198214"

188...541

/locus_tag="SF2947"

188...541

/locus_tag="SF2947"

/notes="Residues 1 to 114 of 117 are 80 pct identical to

pneumoniae gb: AAF06334.1"

/codon_start=1

/transl_table=11

/product="orf, partial conserved hypothetical protein"

/protein_id="AAN44428.1"

/db_xref="GI:24053359"

/translation="MKRSFTSLVCSLLALVSTGVNAEKVALKLAHNTLERSHVH

QSFELAKEVKSQGNVIRYPSSQGNARETWELLQALDGMTKGSASDLESFDN

IYAIYNLPILSCFW"

558...1151

/locus_tag="SF2948"

558...1151

/locus_tag="SF2948"

/note="Residues 2 to 197 of 197 are 48 pct identical to

residues 145 to 341 of a 341 aa protein from *Bacillus*

halodurans ref: NP_241567.1"

/codon_start=1

/transl_table=11

/product="orf, partial conserved hypothetical protein"

/protein_id="AAN44429.1"

/db_xref="GI:24053360"

/translation="MDSTKDKGFALSAVAGTTSFYAKPITKPEDLKLKIRVQPS

PTTKMELMGSPPTPLISFEVVTAMQGVVDGANNVPSWQTRHIEIAKVFSEDEH

ASIPDELAIVSTKWNKLTPEQQILETAARKSEAYQQKLEKIDADTRAQAKAMGBI

VYDKAPRAVQPLFDFFKPKQQAALLEKFDNAQ"

1352...1699

/gene="dctQ"

/locus_tag="SF2949"

1352...1699

/gene="dctQ"

/locus_tag="SF2949"

/note="Residues 1 to 108 of 115 are 75 pct identical to

residues 22 to 129 of a 136 aa protein from *Klebsiella*

pneumoniae gb: AAF01413.1"

/codon_start=1

/transl_table=11

/product="putative small integral C4-dicarboxylate

membrane transport protein DctQ"

/protein_id="AAN44430.1"

/db_xref="GI:24053361"

/translation="MMIWIWGLLGAAYTVGAQRHLSIDLPAALNKRKQLLSIVINVL

LNQATVSGSDSDIYTVWQADYELLEQTWTDKPKVFSAMLSKGTSPASGWS

VNFYSFQAASDRGRVDDIKTNKYLIVNSDFNYRFSQLESALNTQKNSIPALEKE

gene

ILGRAGSVIVTGGLKLDIKTLATSOVSAAOMQIPMGYVYIILPLSGVMFMFALCFINQ

SIQQLKQPAQAS"

complement (1722. .2117)

/gene="insB"

/locus_tag="SF2950"

complement (1722. .2117)

/gene="insB"

/locus_tag="SF2950"

/note="Residues 1 to 131 of 131 are 100 pct identical to

residues 37 to 167 of a 167 aa protein from *Shigella*

flexneri ref: NP_085393.1"

/codon_start=1

/transl_table=11

/product="orf, conserved hypothetical protein"

/protein_id="AAN44431.1"

/db_xref="GI:24053362"

/translation="MIVCAEMDEHWGVGAKSRQWLFIYAYDIRRIRRVVAHVFGERTL

ATUERLLSLLSAFEVVMVMTDGMPLYESRLKGLHVISKRYTORIERHNLNRQLHAR

LGRKSLFSKSVELDKVIGHYLNKHQ"

complement (2144. .2419)

/gene="insA"

/locus_tag="SF2951"

complement (2144. .2419)

/gene="insA"

/locus_tag="SF2951"

/note="Residues 1 to 91 of 91 are 98 pct identical to

residues 1 to 91 of a 91 aa protein from *Shigella flexneri*

ref: NP_085394.1"

/codon_start=1

/transl_table=11

/product="orf, conserved hypothetical protein"

/protein_id="AAN44432.1"

/db_xref="GI:24053363"

/translation="MASISIRCPSCSATEGVVRNGKSTAGHORVLCSHCRKTWQLOFT

YTAQPGTHQKIIDMAMNGVGCASRIMVGLNVLRLHKNKSGSR"

2578...3777

/locus_tag="SF2952"

2578...3777

/locus_tag="SF2952"

/note="Residues 1 to 395 of 399 are 61 pct identical to

residues 38 to 431 of a 435 aa protein from *Salmonella*

typhimurium LT2 ref: NP_462086.1"

/codon_start=1

/transl_table=11

/product="putative integral membrane protein, possible

transporter"

/protein_id="AAN44433.1"

/db_xref="GI:24053364"

/translation="MLLMFPWDIAAITVSORLANGLONFALLAIPFIFAGTILMNSGG

IAIRLINLAQVMVGRVPSGLGHVNVLANMMEGSLGSAVAANAAGVGLNPIOTKEGY

DPAFTAVNVSSCITGLIIPSNVLIIVSLTAGVSVASLPMAGYLPGLMGLAMIV

CGIIAKRGVPLSERATPAQAKAFDALPSKALVLLVFMGGILGFIATEASAIAVV

YRILSVLIYREVKNRHLPLILLESVVTTISVILLGIFGVSFGMSWAMTADIPYMISA

LMGISNPLILILINIVLIVGIFMDIAPVILFIFPLIAQELGMDPVHFGIMMV

ANUCIGLLTPPVGSALFVSGISGVKIQHLIKELLLFYAALLVALLMVTITPQISLFI

PQLLGLM"

complement (3878. .4804)

/gene="yggM"

/locus_tag="SF2953"

complement (3878. .4804)

/gene="yggM"

/locus_tag="SF2953"

/note="Residues 1 to 308 of 308 are 99 pct identical to

residues 13 to 320 of a 335 aa protein from *Escherichia*

coli K12 ref: NP_417431.1"

/codon_start=1

/transl_table=11

/product="putative alpha helix chain"

/protein_id="AAN44434.1"

/db_xref="GI:24053365"

/translation="MLMTGNWADGEPPTENILKQFKQYHGLKLDALSKNLDAK

GNQATVSGSDSDIYTVWQADYELLEQTWTDKPKVFSAMLSKGTSPASGWS

VNFYSFQAASDRGRVDDIKTNKYLIVNSDFNYRFSQLESALNTQKNSIPALEKE

gene

ILGRAGSVIVTGGLKLDIKTLATSOVSAAOMQIPMGYVYIILPLSGVMFMFALCFINQ

SIQQLKQPAQAS"

complement (1722. .2117)

/gene="insB"

/locus_tag="SF2950"

complement (1722. .2117)

/gene="insB"

/locus_tag="SF2950"

/note="Residues 1 to 131 of 131 are 100 pct identical to

residues 37 to 167 of a 167 aa protein from *Shigella*

flexneri ref: NP_085393.1"

/codon_start=1

/transl_table=11

/product="orf, conserved hypothetical protein"

/protein_id="AAN44431.1"

/db_xref="GI:24053362"

/translation="MIVCAEMDEHWGVGAKSRQWLFIYAYDIRRIRRVVAHVFGERTL

ATUERLLSLLSAFEVVMVMTDGMPLYESRLKGLHVISKRYTORIERHNLNRQLHAR

LGRKSLFSKSVELDKVIGHYLNKHQ"

complement (2144. .2419)

/gene="insA"

/locus_tag="SF2951"

complement (2144. .2419)

/gene="insA"

/locus_tag="SF2951"

/note="Residues 1 to 91 of 91 are 98 pct identical to

residues 1 to 91 of a 91 aa protein from *Shigella flexneri*

ref: NP_085394.1"

/codon_start=1

/transl_table=11

/product="orf, conserved hypothetical protein"

/protein_id="AAN44432.1"

/db_xref="GI:24053363"

/translation="MASISIRCPSCSATEGVVRNGKSTAGHORVLCSHCRKTWQLOFT

YTAQPGTHQKIIDMAMNGVGCASRIMVGLNVLRLHKNKSGSR"

2578...3777

/locus_tag="SF2952"

2578...3777

/locus_tag="SF2952"

/note="Residues 1 to 395 of 399 are 61 pct identical to

residues 38 to 431 of a 435 aa protein from *Salmonella*

typhimurium LT2 ref: NP_462086.1"

/codon_start=1

/transl_table=11

/product="putative integral membrane protein, possible

transporter"

/protein_id="AAN44433.1"

/db_xref="GI:24053364"

/translation="MLLMFPWDIAAITVSORLANGLONFALLAIPFIFAGTILMNSGG

IAIRLINLAQVMVGRVPSGLGHVNVLANMMEGSLGSAVAANAAGVGLNPIOTKEGY

DPAFTAVNVSSCITGLIIPSNVLIIVSLTAGVSVASLPMAGYLPGLMGLAMIV

CGIIAKRGVPLSERATPAQAKAFDALPSKALVLLVFMGGILGFIATEASAIAVV

YRILSVLIYREVKNRHLPLILLESVVTTISVILLGIFGVSFGMSWAMTADIPYMISA

LMGISNPLILILINIVLIVGIFMDIAPVILFIFPLIAQELGMDPVHFGIMMV

ANUCIGLLTPPVGSALFVSGISGVKIQHLIKELLLFYAALLVALLMVTITPQISLFI

PQLLGLM"

complement (3878. .4804)

/gene="yggM"

/locus_tag="SF2953"

complement (3878. .4804)

/gene="yggM"

/locus_tag="SF2953"

/note="Residues 1 to 308 of 308 are 99 pct identical to

residues 13 to 320 of a 335 aa protein from *Escherichia*

coli K12 ref: NP_417431.1"

/codon_start=1

/transl_table=11

/product="putative alpha helix chain"

/protein_id="AAN44434.1"

/db_xref="GI:24053365"

/translation="MLMTGNWADGEPPTENILKQFKQYHGLKLDALSKNLDAK

GNQATVSGSDSDIYTVWQADYELLEQTWTDKPKVFSAMLSKGTSPASGWS

VNFYSFQAASDRGRVDDIKTNKYLIVNSDFNYRFSQLESALNTQKNSIPALEKE

Location/Qualifiers
1 292805

```

/protein_id="AAP18147.1"
/db_xref="GI:30042423"
/translation="MLLVGSLILQGMQSDRSFASRVSMESQSLRQAIVQSALAWG
KXHSWQTQPAVQCSQVAGTDAQVCLRLADNEALLIAGVEGSLWRTGEVTDGKIVFS
PRGWSDFCLKEGALCOLP"
complement(4239..4802)
/gene="ppdB"
/locus_tag="S3033"
complement(4239..4802)
/gene="ppdB"
/locus_tag="S3033"
/function="putative enzyme; Not classified"
/notes="residues 1 to 187 of 187 are 98.39 pct identical to
residues 1 to 187 of 187 from Escherichia coli K-12 :
B2825"
/codon_start=1
/transl_table=11
/product="prepilin peptidase dependent protein B"
/protein_id="AAP18148.1"
/db_xref="GI:30042424"
/translation="MPVKEOGFSLEVLIAAISVLLGNARFLPALORESLSTSTRK
LAEDEIWLRFVTVAKHLORAGYCHSGTCEGLIEIVGDCIIVQWDANSNGIWDREP
VKESDQIGRLKEHVLFTLAGATSCRGKGMKVTPDAIIIDTFQVVRQDVSGFSPL
TVNWRASKSEPQVVDASVSTGFNL"
complement(4793..5263)
/gene="ppdA"
/locus_tag="S3034"
/notes="point mutation"
/pseudo
complement(5447..6241)
/gene="thyA"
/locus_tag="S3035"
complement(5447..6241)
/gene="thyA"
/locus_tag="S3035"
/function="enzyme; 2-Deoxyribonucleotide metabolism"
/notes="residues 1 to 264 of 264 are 99.62 pct identical to
residues 1 to 264 of 264 from Escherichia coli K-12 :
B2827"
/codon_start=1
/transl_table=11
/product="thymidylate synthetase"
/protein_id="AAP18149.1"
/db_xref="GI:30042425"
/translation="MKQYLELMQKVLDEGQTKNDRTGTGTLISFGHOMRENLQDGFPL
VTKRCHRSIIHELLFQGTNTLAIHENNVITWDEWDENGDLGVYKQGRWMP
TPDGRHQITVTLNQLKNDPSRRIIVSAWNVGLDKALAPCHAFFQFYVADGKLS
COLYRSCDVFLEGTASVALLVHMAQQDILEVGFVMTGDTLHSHNMQDTHL
QLSRPRLPLKLIIRKRPESIFDYPFEFELEGDPHGVKAPVAI"
complement(6248..7123)
/gene="lgt"
/locus_tag="S3036"
complement(6248..7123)
/gene="lgt"
/locus_tag="S3036"
/function="enzyme; Macromolecule synthesis, modification:
phospholipids"
/notes="residues 1 to 291 of 291 are 95.87 pct identical to
residues 1 to 291 of 291 from Escherichia coli K-12 :
B2828"
/codon_start=1
/transl_table=11
/product="phosphatidylglycerol-prolipoprotein
diacylglyceryl transferase; a major membrane phospholipid"
/protein_id="AAP18150.1"
/db_xref="GI:30042426"
/translation="MPSVYLHPFEPDPVTFISGFVALHWYGLMYLVCFFIAPMWLATRR
ANRPGSGWTKNEVNELLYAGFLVGGGRIGYVLFYFNFFQFMADPLFLFRVMDGMSF
HGSLGIVVMILIFARKTKRSFFQSDFTAPLIFGLGAGRLCNFLNGELWGRVDNF
PFWLPLFGRTDILLQNTQWQSFTDFTYGLPRHPSQLYELLLEGVLFITLNIYI
RKRPMPWAGSLGLICGAFRIIIVEFRQDAQFTGAWVQYISMGQILSIPMIVAGVI
MMWVYRRSPQOHVS"
complement(7274..9520)

```

Db 120391 ACACGAGGATAGCCATTCGATGCTCTAAGCTAATGAGCTGCCGAAAGTCGGCATCG 120332
 QY 779 TTTACGCTCACCAGATGATACGATGTTTATGCTCAATGACGCTTTCAGGAGGAGCA 838
 Db 120331 TTTTAACTACGCTAAGCAGATCCGATCTTCGCGCTAAAGCAGCTGTAGATGCGGCTATG 120272
 QY 839 AAGGAATCATCCATGAGGAGGATGCGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTC 898
 Db 120271 ATGGCATCTGTTAGCGCTGGTGTGGTAAATGTAACCTGTATAAATCGTGTTCGACACCC 120212
 QY 899 TTGAAAGAGCAGCAATCATGAGGCTAGTCTCGCTCGAAGCTCTAGATGGGAGTGTGTT 958
 Db 120211 TGGCAACCGCGCGGAGAAACGCGACTGCGATGAGTGGCTTCTCCGCGTACCGAGCGGTG 120152
 QY 959 CCACCCAGGAGGCTGGAAGTGAATGAAGAACTTGGTTTGTGGCTACAGAGATC 1018
 Db 120151 CTACCATCAGGATGCTGAAGTGAATGATGCGAATACGCTTCTCGCTTGCACGC 120092
 QY 1019 TCAACCTCAAAAGCAGAGTGTCTTATGTTAGCCCTCACCAAACTAGTATGATAG 1078
 Db 120091 TGAACCGGAGAAAGCGCGCTGCTGCTGAGCTGGCTCTGACGCAAAACCAAGATCCG 120032
 QY 1079 AGGCGATCAAAAGATCTTCTCCACCTATTATC 1112
 Db 120031 AGCAGATCCAGCAGATTTTCAATCAGTACTAATC 119998

RESULT 9
 AE005526/c
 LOCUS
 DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 145 of 290.
 ACCESSION AE005526 AE005174
 VERSION AE005526.1 GI:12517505
 KEYWORDS
 SOURCE
 ORGANISM Escherichia coli O157:H7 EDL933
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 1 (bases 1 to 10749)
 Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.P., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.
 Genomic sequence of enterohaemorrhagic Escherichia coli O157:H7
 Nature 409 (6819), 529-533 (2001)
 21074935
 11206551
 2 (bases 1 to 10749)
 Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.P., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.
 Direct Submission
 Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 Location/Qualifiers
 1. .10749
 /organism="Escherichia coli O157:H7 EDL933"
 /mol_type="genomic DNA"
 /strain="EDL933"
 /serotype="O157:H7"
 /db_xref="taxon:155864"
 /note="enterohemorrhagic complement (89. .1135)"
 /gene="ansB"
 /note="synonym: 24302"
 complement (89. .1135)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 gene
 CDS

Query Match 26.5%; Score 300; DB 1; Length 292906;
 Best Local Similarity 57.1%; Pred. No. 1.6e-74;
 Matches 568; Conservative 0; Mismatches 420; Indels 6; Gaps 1;
 QY 119 TGGCTAAACCCCAAGTACTCTCTGAGGAGCTCACCGTTGATAGCTTCTTGACGCG 238
 Db 120985 TGGCAATACCAATATCACCATTATTTAGCAACCGCGGACCACTTCGCGTGTGTGACT 120926
 QY 179 CTAGCGTCAAGTAGTACTCTCTGAGGAGCTCACCGTTGATAGCTTCTTGACGCG 238
 Db 120925 CGCAACCAATCTAATACACACGCGGTAAAGTTGGGTAGAAAATCTGTTAATCGG 120866
 QY 239 TCCCTGCCATCAACGACCTAGCACCATCAAGGGTGAACAGATCTCAAGCATTTGCTCC 298
 Db 120865 TGCCGCAACTAAAGACATTTGCAACCTTAAAGGCGAGCAGGTAGTGAATATCGCTCC 120806
 QY 299 AAGATGACGGGTAAAGTGTGGCTTAACTAGCAGAGCTGCTCAATGAGTCTCTCGCC 358
 Db 120805 AGGACATGAACGATAATGCTGGCTGACACTGGCGAATAAATTTAA-----CGCGACT 120752
 QY 359 AAAAGAGACCGGCGGTGATCATCACCCATGGAACTGACACCATGGAGAGACGGT 418
 Db 120751 GCGATAAACCGACGGCTTCGTCATTTACCCAGGTACCGACGATGGAAGAACTGCTT 120692
 QY 419 TCTTCTCACTCAGCGTGAAGAAACCAAAACCTGCTGCTGCTGAGCGGCATCGCTC 478
 Db 120691 ACTTCTCAGCTGACGGTGAATGCGAATGCGAACCCTGCTGCTGCTGCTGCTGCTG 120632
 QY 479 CAGGCTCTTCATGAGTGTGATGGCCCATGATCTCTATACCGCGGTGAATGTAGCA 538
 Db 120631 CGTCCAGTCCATGAGCGAGACGGTCTCATCTCACTGTATACGCGGTAGTACCGCAG 120572
 QY 539 TCAACAAAGCCTCTACTAACAAGGAGTGTGATGATGATGAACGATGATTCACGCCG 598
 Db 120571 CTGATAAGCTCCGCTAATCGTGGCGTGTGCTGCTGATGAACGACACCGCTACTGACG 120512
 QY 599 CCAGAGAGGACCAAGCTCAACACCCAGCTCAATGATTTGCTTCGCCCAACACAG 658
 Db 120511 GTCGCGATGTACCAAAACCAACCAACCGACGTAGCGCTTCAAGTCTGTTAATACG 120452
 QY 659 GTAATTCGCGACAGTCTATTATGGCAAGTTCGAGTATTTCACTCAATTCGCTTCGACCTC 718
 Db 120451 GTCTCTGGGATACATTCACACCGTGAATGATTTGATACCAACGTTACCCGCGCAGTAAGC 120392
 QY 719 ACACCTTGAAGTGAAGTTTGTATATTAGCAAAATCGAAGAACTCCCGCAGATCGATTC 778

/gene="ptsp"
 /locus tag="S3037"
 complement (7274. .9520)
 /gene="ptsp"
 /locus tag="S3037"
 /functions="enzyme; Transport of small molecules: Other"
 /note="residues 1 to 736 of 748 are 99.59 pct identical to residues 1 to 736 of 748 from Escherichia coli K-12 ; B2829"
 /codon_start=1
 /transl_table=11
 /product="PTS system, enzyme 1, transcriptional regulator (with NFR and NFR proteins)"
 /protein_id="AAP18151.1"
 /db_xref="GI:30042427"
 /translation="MTRLEIVEKVASAPRLNEALNIVLDICLAWDEVCVYLAD HDRCYVLMATRGKLRGTVTLAFDEGIVGLVRLAEPLINLADAKHPSFYIIPSV KEERFALGVPIIQRELGLVAVQORELROYDESESLVTLATQMAALRSQSQT ALFGQYQTRIRALPAAPGVAIAEGMDATLPLMEOVYQASTLDPALERLTLGALBE AANEFYRSKFRPAGAKETAALFDLXSHLLSDRLRLRELFADVDKGSVAEMAVKTVI EKFAEQFALSDNYLKRAGDLRALGRLIFLHDDANOGNENAWPERFILVADELSAT LLAELPQRLVGVVRDGAANSHAAIMVRALGIPVMGADIQPSVLHRRRTLIVSYRGE LLLDPQVLLQEVYORLISEIEISRLAEEDVNLPAQKSGERIKWMLNAGLSPEHEK LGSCIDGIGLYRTIIPMLQSGFPSEEOVAQYQMLQMFNDKPVTLRLDVGADKOL PYMPISENPCLHWRGIRITLDQPEIFLIQVRLMRLANATGNLILPMVTSLEVD EARLIERAGREVEIMIGIEIPRIGIMLEVPSPVPMPLHAKRVDVFISVGTNDLQ YTLAVDRNTRVANIYDSLHPMLRALAMTAREAEIHGIDILRLCGEMAGDPMCAVILI

```

/gene="ansB"
/function="enzyme; Degradation of small molecules: Amino acids"
/note="Residues 1 to 348 of 348 are 98.85 pct identical to residues 1 to 348 of 348 from Escherichia coli K-12 Strain MG1655: B2957"
/codon_start=1
/transl_table=11
/product="periplasmic L-asparaginase II"
/protein_id="AAG58088.1"
/db_xref="GI:12517508"
/translation="MEFFKKTALAALVMGFGAALALPNITILATGTTAGGDSATK SYTAGKGVENLVAAPQLDIANVKGQVNNIGSQMNDVWLTLAKINTDCDT DGFVTHGTDTEETAYFLDLTKCDKPVVMVGAMPSTMSADGPNLYXAVTAAAD KASANRGLVNVNDVLDGRDVTNTKTNDVATFKSVNGPLGYHNGKIDVQRTAPARK HTSDTDFVDSKLNELPKVGIYVYANASDLPKALVDAGYDGVISAGVNGNLYKSVF DTLATAKNGTAVRVSRRVPTGATTQDAEVDADKYGFATSGTLNPKQARVLLQLALTO TKDPOIQIIFNOY"
/complement(13111..2030)
/gene="yggN"
/note="synonym: Z4303"
/complement(13111..2030)
/gene="yggN"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2958"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58089.1"
/db_xref="GI:12517507"
/translation="MMRMLLAALSVTAMTAHADYQCSVTPRDDVIVSPQTVKGS NGNLVITPDGVMYNGQYSLNAAQEOAKDYQAEILRSLTPWIDEGAKSVKARIAL DKIVQMGSSSKMRSLTKLDAOLKQOMNRIETRSDDLTEHYKALDQVXEQQLV NOAMGILQDSINEMGAKAVLKSGNLQNVLSGLQSSIQIEWKKQKNDPOFGK DVCGRVTVLDSRKALVGNLK"
/complement(2214..2570)
/gene="yggL"
/note="synonym: Z4304"
/complement(2214..2570)
/gene="yggL"
/function="orf; Unknown function"
/note="Residues 1 to 118 of 118 are 99.15 pct identical to residues 1 to 118 of 118 from Escherichia coli K-12 Strain MG1655: B2959"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58090.1"
/db_xref="GI:12517508"
/translation="MGLNVREGIEMAKNRRLRRKKVHIDFQELGFSVAMRPFQSTS BEQIDKTVDDFINEVIEPNKLAEPDGSYLAWGLICQWETGKCTEHBQAIVRKLEER NLDEVRTSELDVYWD"
/complement(2540..3259)
/gene="yggH"
/note="synonym: Z4305"
/complement(2540..3259)
/gene="yggH"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2960"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58091.1"
/db_xref="GI:12517509"
/translation="MKNDVISPFDENGRLRIRSVRQGLTKGQHALENTWVP MGVTSEMDLDPALFGRAPVTLTGFGMGASLVAMAKREQDFLGLIEVHSPGVGA CIASAHEGLSLNRYMCHDAVNLHKIIPDNSLRMVQVFFPDPWHRKRRIRVQVP FAEALYKSLQGLGVFHMTADWEPFAEHMIEVMSSIDGYKNLSESDYVPRASRPVTK FEQRHRLGHGVNDLWFERVK"

```

```

gene
CDS

```

```

3420..4472
/gene="mutY"
/note="synonym: Z4306"
3420..4472
/gene="mutY"
/function="enzyme; Macromolecule synthesis, modification: DNA - replication, repair, restr./modific'n"
/note="Residues 1 to 350 of 350 are 99.42 pct identical to residues 1 to 350 of 350 from Escherichia coli K-12 Strain MG1655: B2961"
/codon_start=1
/transl_table=11
/product="adenine glycosylase; G.C --> T.A transversions"
/protein_id="AAG58092.1"
/db_xref="GI:12517510"
/translation="MQASQFSAQDLVDYKGRKTLPMQIDKTPVKVLSVWLQOQTQ VATVPIPERFEMARFPTVTLANAPLOEVLHLWTGLGYARARNLHKAQQVATLHGG KPEPTEFEEVALPGVSTAGAILSLSGKHFFPLDGNVKKVRLAKYAVSGWPGKKEV ENKLWSLSEQVTPAVGVRFNQAMDLGAMICTRSKPKSLCPLQNGCIAATNNSWSL YPGKKPKQTLPERTGYFLLQDEHDEVLIAQRPSPGLWGGLYCFQFQADESLRWLAQ RQIAADNLTLQTLAFRHTFHFHLDIVPMWLPVSSFTGCMDEGNALWYNLAQPPSVGLA APVERLLQQLRTGAPV"
4500..4775
/gene="yggX"
/note="synonym: Z4307"
4500..4775
/gene="yggX"
/function="orf; Unknown function"
/note="Residues 1 to 91 of 91 are 100.00 pct identical to residues 1 to 91 of 91 from Escherichia coli K-12 Strain MG1655: B2962"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58093.1"
/db_xref="GI:12517511"
/translation="MSRTICTFLQREAGQDPQLYPGELGKRIYNEISKEAWAQWH KOTMLNEKKLANMNAEHRKLLEQEMVFLFEGKEVHIEGTYTPEDKK"
4837..5919
/gene="mltC"
/note="synonym: Z4308"
4837..5919
/gene="mltC"
/function="putative enzyme; Cell envelop: Murein sacculus, peptidoglycan"
/note="Residues 1 to 360 of 360 are 99.72 pct identical to residues 1 to 360 of 360 from Escherichia coli K-12 Strain MG1655: B2963"
/codon_start=1
/transl_table=11
/product="membrane-bound lytic murein transglycosylase C"
/protein_id="AAG58094.1"
/db_xref="GI:12517512"
/translation="MMKYLALALAPLISCTTKKGTNEAWKDTNGFDILMQQ FAHNTEIWGFKEVYAGPKYKVTQYTRSHINFDGTTITETIAGTEPAHLRR AITILLMGDDSPSDLYSDVDITISKEPFLYGVQVNTGPIRWEGRASNFADYLL KNRLKRSNGLRIYVVTINMVDLKAHKLGMVROASKYGVDESLILAIMQTE SSFNPYAVSRSDALGMQVQHTAGKDVRSOGKSGTSPSRFLDPASNDITGTAYLA MLNNVILGIDNPTSRRTAVITAYNGAGSVLRVFSNDKIQAAANIINTWTPGDDVQTL TRHPSASRRYLYKVNTAQSRYRR"
6073..7377
/gene="nupG"
/note="synonym: Z4309"
6073..7377
/gene="nupG"
/function="transport; Transport of small molecules: Nucleosides, purines, pyrimidines"
/note="Residues 1 to 434 of 434 are 99.76 pct identical to residues 1 to 434 of 434 from Escherichia coli K-12 Strain MG1655: B2964"
/codon_start=1
/transl_table=11
/product="transport of nucleosides, permease protein"

```

```

gene
CDS

```


/protein_id="AAG58095.1"
/db_xref="GI:12517513"
/translation="MCFKHSASRCEENNNKLQKILSLFQCLWGLWTLTGLSY
MFVTLKFDGASIGAVYSLGIAAFAVPMALLIGIVADKWLAKVVAICHTIGATLFWA
AQVTPPAMELVILINSFAYMPTGLINTISYVRLQAGMDIVDRPPIRWITGFIGI
MAMVVSLSGELSHMQLYIGALSAIYLFTLPHIPVAKQOANOSWTLGLGLDAP
ALFKKEMALFFIFSMILGABQLTNMFGNTFLSKDKMPFASFIQVHASIIMSIS
QISFTLILTFPFLSYRIGNVMISIVAMIRFALFYAGDTPFGTLLVLLSMIVY
GCAPDFNIGSVFVEKEVSPAIRAAQGMFLMNTNGFCILGIVSGKVMYMTONG
ITDQWTLIPAGYSVVLAFAPMAMFKYKVRVPTGTQTVSH"
complement (7427..9622)

Query Match	26.2%; Score 296.8; DB 1; Length 10749;
Best Local Similarity	56.9%; Pred. NO. 8.1e-74;
Matches 566; Conservative	0; Mismatches 422; Indels 6; Gaps 1;
QY	119 TGGCTTAACCCCAAGTGACTATCTTAGCCACAGGAGGACCATCGCTGGTTCGGGGGAAT 178
DB	1074 TGGCATTAACCAATATACCAATTTAGCAACCGGGGACCATTCGCCGGTGGTGTGACT 1015
QY	179 CTAGCGTCAAGAGTAGTACTCTGTGTGAGCAGTACCGTGTGATAGCTTCTTGACGGC 238
DB	1014 CGCAACCAATCTAACTACACAGCGGTAAAGTTGGCGTAGAAAATCTGGTTAATGCGG 955
QY	239 TCCCTGCCATCAACGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCC 298
DB	954 TGCCGCAACTGAAGACATTTGGAACGTTAAAGCGGAGCAGGTAGTGAATATCGGCTCCC 895
QY	299 AAGAGATACGGGTAAAGTGTGGCTTAACTAGCAAGCGTGTCAATGAGTCTCTCGGCC 358
DB	894 AGGACATGACGATGTCTGCTGACACTGGCGGAAAAAATTAAC-----CACCGACT 841
QY	359 AAAAAAGACGACGCGGTGATCATCACCCATGGAACCTGACACCATGAGAGAGCGGCTT 418
DB	840 GCGATAAAACCGACGCTTTCGTCATCTACCGGTACCGACATGGAAGAAACCGGCTT 781
QY	419 TCTTCTCAACTCAGCGTGAAGAAACGCAAAACCTGTCTGCTGTGAGCGGCATCGCGTC 478
DB	780 ACTTCTCGACTGACGTGGAATGCGCAACCGGTGTGATGTTGCGCGCAATGCGTC 721
QY	479 CAGGCTCTTCCATGAGTGTGATGCGCCCATGAATCTCTATACGCGGTGAATGTAGCGA 538
DB	720 CGTCCAGCTTATGAGCGCAGCGGTCCATTCAACCTGTATACGCGGTAGTGACCGCAG 661
QY	539 TCACAAACCTCTACTACAAAGAGTGGTGTGATGATGACGATGATGATTCACGCGC 598
DB	660 CTGATAAAGCTCCGCTTATTCGTTGGCGTGTGTTGATGATGACGACACCGTACTGAGC 601
QY	599 CCAGAAAGCGACCAAGCTCAACACCGCGCAGTCAATGCTTGTCTCGCCCAACACAG 658
DB	600 GTCCGATGTACCAAAACCAACACCGACGATGACGACCTTCAAGTCTGTTAACTAG 541
QY	659 GTAAATCGGCACAGTCTATTATGGCAAGTCGAGTATTTCACTCAATCCGTTGACCTC 718
DB	540 GTCCGCTGGGATACATTCACACGGAAGATTGACTACCAACGTACCCCGGACGTAAGC 481
QY	719 ACACCTCTGCAAGTGATTTGATTTAGCAAAATCGAAGAACTCCCGACGATGCGATATTC 778
DB	480 ACACGAGGACACCGCGTTCGATGCTCTTACGCTGATGATGATGATGATGATGATGAT 421
QY	779 TTTACGCTCAACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
DB	420 TTTTAACTAGCTACGCGATCCGATCTTCCGCTAAAGCCCTGGTAGATGCGGGCTATG 361
QY	839 AAGGAATCATCATGACGAGCATGGGCAATGGGACCTTTTCCCTTTTGACTCAAAATGCTC 898
DB	360 ATGGCATCTGAGCGCTGCTGGTAAACCGGCAACCTGATATAATCCGTTGTTGACACGC 301
QY	899 TTGAAAAGCAGCAAAATCAGGCGTGTCTGCTCGCTCGAAGCTCTAGAGTGGCGAGTGGTT 958
DB	300 TGGGACCGCGGGAAGAAACGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
QY	959 CCACCAACCAAGAGGCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018

DB	240 CTACCACTCAGGATCGCGAAGTGGATGATGCGAATACGGTTTATTGCTCTGCTAGCG 181
QY	1019 TCACCCCTAAAAGCCAGAGTGGTCTTATGTTAGCTTACCAAACTAGTATGATAG 1078
DB	180 TGAACCCGCAAAAAGCGCGGCTCTGCTCAACTGGCTGACGCAACCAAGATCCGC 121
QY	1079 AGGCGATCCAAAGATCTTCTCCACCTATTAAATC 1112
DB	120 AGCAGATCCAGCAGATCTTCAATCAGTACTAATC 87

RESULT 10	AP002563/c
LOCUS	266658 bp DNA linear BCT 07-MAR-2001
DEFINITION	Escherichia coli O157:H7 DNA, complete genome, section 14/20.
ACCESSION	AP002563 BA000007
VERSION	AP002563.1 GI:13363121
KEYWORDS	Escherichia coli O157:H7
SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
ORGANISM	1 (sites)
REFERENCE	1 (sites)
AUTHORS	Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.
TITLE	Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
JOURNAL	Genes Genet. Syst. 74 (5), 227-239 (1999)
MEDLINE	20198780
PUBMED	10734605
REFERENCE	2 (sites)
AUTHORS	Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and Hayashi, T.
TITLE	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655
JOURNAL	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
MEDLINE	20557356
PUBMED	11108008
REFERENCE	3 (sites)
AUTHORS	Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.
TITLE	Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak
JOURNAL	Gene 258 (1-2), 127-139 (2000)
MEDLINE	20564182
PUBMED	11111050
REFERENCE	4 (sites)
AUTHORS	Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and Shinagawa, H.
TITLE	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12
JOURNAL	DNA Res. 8 (1), 11-22 (2001)
MEDLINE	21156231
PUBMED	11258796
REFERENCE	5 (bases 1 to 266658)
AUTHORS	Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and Hayashi, T.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:ken@gen-info.osaka-u.ac.jp)

URL: <http://www.gen-info.osaka-u.ac.jp/>, Tel: 81-6-6879-8365,
Fax: 81-6-6879-2047

COMMENT
FEATURES

Genome project.

Location/Qualifiers

1. .266658

/organism="Escherichia coli O157:H7"

/mol_type="genomic DNA"

/strain="O157:H7"

/sub_strain="RIMD 0509952"

/db_xref="taxon:83334"

complement(244. .693)

/gene="ECs3650"

complement(244. .693)

/gene="ECs3650"

/note="similar to B2790 ECOLI gi|1789154 percent identity

100 in 149 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37073.1"

/db_xref="GI:13363122"

/translation="MAEIGIFVTWYGNLSLVAEBAELLTAQGHKATVPEDPEISDW

LPYODKYVLVVTSTTGGDLPDSIVPLFCGIDKSLGQPNLRYGVIALGDSYVNFPCN

GKQFDALLQSQARVGMELLDIASENPEPESNPVWEQMGTLSS"

complement(711. .1493)

/gene="ECs3651"

complement(711. .1493)

/gene="ECs3651"

/note="similar to YQCH ECOLI gi|1789155 percent identity

99 in 260 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37074.1"

/db_xref="GI:13363123"

/translation="WLEILYODEWLVAVNKPQGMVLRHSWLDREKVVVMQTVRQIG

QHVTAHRLDRPTSGVLLMGLSSAGRLLAQOFQHQIQKRYHAIVRWGLMBEAVLDY

FLVEELDKADKFAREQPQPVAVHYRGLATVEMPTATGYPTTRYGLVLEPEKTGR

KHQRHLAHLRHPIDGSKHGDLRQNRSGAEHFGQLRLMLHASQLSLTHPTGTEPLT

IHAGLDITWQALSQFGWRGLPENRVEFSAPSQDGERSS"

complement(1493. .1822)

/gene="ECs3652"

complement(1493. .1822)

/gene="ECs3652"

/note="similar to B2792 ECOLI gi|1789156 percent identity

98 in 109 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37075.1"

/db_xref="GI:13363124"

/translation="WTHDRVRLQLOALEALLRHOHWNRDEPOPHQFNSTQPFMDT

MEPLEWLOWLIPRMHDLNNQLPGFAVAPYVEMALTDHPQRLIALLEKLDA

LPADSDAS"

complement(1865. .2233)

/gene="csrB"

complement(1865. .2233)

/gene="csrB"

/product="CsR RNA"

/note="inhibitor of CsR"

complement(2444. .2989)

/gene="ECs3653"

complement(2444. .2989)

/gene="ECs3653"

/note="similar to SYD ECOLI gi|1789157 percent identity 98

in 181 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="Syd protein"

/protein_id="BAB37076.1"

/db_xref="GI:13363125"

/translation="MDDLTAQAKDFTARYCDAMWEEHKSWPLSBEYGVPSPCIIST

TEDAVYWPQPFTEQNNAVERAFDIVIOPTINTFTYTOAGDMHAQFGDIKLTLLQ

TWSEDDPRRQENLIGHLVTKOKRLKLPPTLIATLEEBELVISVCNLSGEVCKETLGT

RKRTHLPANLAEPNLQKPLL"

3057. .3905

/gene="ECs3654"

3057. .3905

/gene="ECs3654"

/note="similar to YQCD ECOLI gi|1789158 percent identity

98 in 282 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37077.1"

/db_xref="GI:13363126"

/translation="MSSYANHOALAGLTCKSTDYRDYDASLLQGVPSLRNRDPLGL

KADNLPQGTDIWTLYELSWLNAGLQVAVGHVELDYTSVNLIESKFSKLYLSNFQ

TRNNDEVRQTLERDLSTCAQGEVSVALYRLDEGQPIGHFNGTCTDDODITIDNY

EFTDYLENATSGEKVVEETLVSHLLSKNCLITHQPDWMSIQIYVRGQIQIDREKLLRY

LVSFRHNHEFCQVERIFENDLLRFCQEKLSVYARYTRGGLDINPWRNSDFFVPT

TRLVRQ"

4017. .5381

/gene="ECs3655"

4017. .5381

/gene="ECs3655"

/note="similar to YGDH ECOLI gi|1789159 percent identity

99 in 454 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37078.1"

/db_xref="GI:13363127"

/translation="MITHSPLGSMDSLEVDMLKRTASSDLYQLFRNCSLAVLNS

GSITNSKELLSRFENFINDVLRERGVKLELNPPEAFVDGRIIRALQANLPAVL

DLFLVYGOJHNTVRPNLNDLSVHTNLVFSILRNARALHVGSEAPNVCWGGSIN

ENEYLXARVGNQLGRLNLCITCGCGAMEAGAAQVHAQKDYKDSRFIGNTEPS

ITAAPVPLVNLLEIIMPDIKRLFAVRIAGHGIIPFGCVGTABEYLLYLLGLMNSA

NKDQPLLLTQKESADYFRLVDFVHTLGENARRHYRIIDDAEAARVAKMKMSMP

LVKENRDTGDAYSFNWSMRIAPDLQMPPEPHENMANLKYPDQPVVLAADLRRAF

SGIVAGNVKEVGIRATEIEFGPYKINGDKIMRMDLLQGFVQAHRMKLPGSAVIPC

ETCT"

5938. .7227

/gene="ECs3656"

5938. .7227

/gene="ECs3656"

/note="similar to SDAC ECOLI gi|1789160 percent identity

100 in 429 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="probable serine transporter"

/protein_id="BAB37079.1"

/db_xref="GI:13363128"

/translation="METTQTSTIASKDSRSARWKTDTMMMLGLYGTAGVGLFLPIN

AGVGMIPLIIMAILAPMTFPAHRLGRFLRVLSKNGPDGDIETVEEHEFGIGAGKLIT

LLYFPAIYPIILLYSVALTINVTVESHSQHGTMTPPPRAILSLILIVGMMTIVRFGQM

IVKMSILVFPFVGVLMLALYLIPOWGALETSLDTSATNGLGLMTLWLAIPVM

VFSFNHSPIISSFAVAKKEEYGDMAEQCKSLFAFIMMVLTMFVFSCVLSLTPA

DLAAAEQNIISLYLANHFNAPVLAWMAPIIAITAITKPLGHYLGAREGFNGMKWK

SLRGKGSIEINKLNRIYALFVLTWTIVATNPISILGMIELTGGPIIAMILLFMPMY

AIOKVPAMRYKSGHISNVFVVMGLIAISFYSLPS"

7285. .8652

/gene="ECs3657"

7285. .8652

Query Match 26.2%; Score 296.8; DB 1; Length 266658;

Best Local Similarity 56.9%; Pred. No. 1.4e-73;

Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

QY	119	TGGCTAAACCCCAAGTGACTATCTTAGCCACAGGAGGCACCATCGCTCGTTCGGGGGAAT	178	ACCESSION	AB016766	AB0114075	
Db	197570			VERSION	AB016766.1	GI:26109707	
				KEYWORDS			
				SOURCE			
QY	179	CTAGCGTCAAGAGTAGTACTCTGTCTGGAGCAGTCAACGGTGTGATTAAGCTTCTTGACGCG	238	ORGANISM	Escherichia coli CFT073		
Db	197510			REFERENCE	Bacteroides; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
				AUTHORS	Enterobacteriaceae; Escherichia		
QY	239	TCCTCTGCATCAAGCAGCTAGCCACCATCAAGGGTGAACAGAGATCTCAAGCATTTGGCTCCC	298		1 (bases 1 to 303121)		
Db	197450			TITLE	Welch, R.A., Burland, V., Plunkett, G.D. III, Redford, P., Roesch, P.,		
				JOURNAL	Rasko, D.A., Buckles, E.L., Liou, S.-R., Zhou, S., Schwartz, D.C.,		
				PUBMED	Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C.,		
QY	299	AAGAGATGACGGGTAGGTGTGGCTTTAAACTAGCCAAAGCGGTCAATGAGCTCTCGGCC	358	REFERENCE	Perna, N.T., Mobley, H.L.T., Donnenberg, M.S. and Blattner, F.R.		
Db	197390			AUTHORS	Extensive Mosaic Structure Revealed by the Complete Genome Sequence		
				TITLE	of Uropathogenic Escherichia coli		
QY	359	AAAGAGACCGGAGCGTGATCATCACCATGGAACTGACACCATGGAAGAGACCGCTT	418	JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)		
Db	197336			PUBMED	12471157		
				REFERENCE	2 (bases 1 to 303121)		
QY	419	TCCTCTCAACCTCACGGTGAAGGCAAAACCTCTCTGCCCTTTAGGCGCATCGCTC	478	AUTHORS	Welch, R.A., Burland, V., Plunkett, G.D. III, Redford, P., Roesch, P.,		
Db	197276			TITLE	Rasko, D.A., Buckles, E.L., Liou, S.-R., Zhou, S., Schwartz, D.C.,		
				JOURNAL	Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C.,		
QY	479	CAGGCTTTCATGAGTGTGATGGCCCCATGAATCTCTATAACGCGGTGAATGTAGCGA	538	REFERENCE	Perna, N.T., Mobley, H.L.T., Donnenberg, M.S. and Blattner, F.R.		
Db	197216			AUTHORS	Direct Submission		
				TITLE	Submitted (20-JUN-2002) Genetics Laboratory, University of		
QY	539	TCACAAAGCTCTACTAACAAAGAGGTGGTGAATTTGTGATGAACAGTAGATTCACGCCG	598	JOURNAL	Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA		
Db	197156			FEATURES	Location/Qualifiers		
				gene	1. 303121		
QY	599	CCAGAGAGCGAACAGCTCAACACACCGGATCAATGATTTGCTTCGCCCAACACAG	658	source	/organism="Escherichia coli CFT073"		
Db	197096			CDS	/mol_type="genomic DNA"		
					/strain="CFT073"		
QY	659	GTAAATCGGCACAGCTATTATCGCAAGTCGAGTATTTCACTCAATCCGTTTCGACCTC	718		/db_xref="taxon:199310"		
Db	197036				177. .1754		
					/gene="ygfU"		
QY	719	ACACCTTTCGAAGTAGTTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTGCATATTC	778		/locus_tag="c3465"		
Db	196976				177. .1754		
					/gene="ygfU"		
QY	779	TTTACGCTACCCCGATGATCTGATGTTTTAGTCAATGATGACAGTCCCGAGAGGCCA	838		/locus_tag="c3466"		
Db	196916				/functions="putative transport"		
					/note="Escherichia coli K-12 ortholog: b2889; Escherichia		
QY	839	AAGGAATCATCATCGAGCATGGCAATGGGAACCCCTTTCCCTTTTGACTCAAAATGCTC	898		coli O157:H7 ortholog: z4226"		
Db	196856				/codon_start=1		
					/transl_table=11		
QY	899	TTGAAAAGACGCCAAATCAGCGGTAGTCGCTCGTTCGAGCTCTAGAGTGGCAGTGTT	958		/product="Putative purine permease ygfU"		
Db	196796				/protein_id="AAN81913.1"		
					/db_xref="GI:26109708"		
QY	959	CCACCAACCAAGAGCTGAAGTGAATGATAAGAACTTTGGTTTGTGGCTACAGAGATC	1018		/translation="WPQFSLTLPKRWHTSMCLDSFMNNGCFLWRKAKETLM		
Db	196736				SAIDSQLPSSGQDRPTDEVDRILSPGKLIILGHVLVMYAGAVPLMIGDRILGLS		
					KEATAMLISSDLFCGIVLQICIGRFGMRILPVMISVTAFTVMTAIGMNPDI		
QY	1019	TCAACCTTCAAAAGCGAGAGTCTTTATGTTAGCCCTCACCACCACTAGTATAGAG	1078		LILGIPGATTAAGTITTLAIPILGRIMPLPELVGTGVTISIGLSIIQVIGDWAAGGK		
Db	196617				NPOYGNPVYLGISFAVLIIFILLITRYAKGFMNSVAVLLGIVFGFLSQMMNEVNLGSI		
					HDASFAVITPMSFGMPTDFPSILITMTAVLIIVFIESMGFLALGETVGRKLSHDI		
QY	1079	AGGCGATCAAAAGATCTTCTCCACCTATTAATC	1112		IRGLRVGVTMIGGTENSFPHTSFQSNQGLVSVTRVHSRWVCITASGIIILFLFGMPK		
Db	196616				MAVLVASIPQFVLGAGLVMEFGVLAATGIRILSRCNVYTNRYNLYIVAILSLGVGMTPT		
					LSHDFPFXKLPVAVLQPLHSGIMLATLSAVLNVFPNGVQHADLVKESVSKDLKVRT		
QY	196616	AGCAGATCCAGCAGATCTTCAATCAGTACTATC	196583		VRMVLMLRKLKKNHEG"		
Db					1738. .1881		
					/locus_tag="c3466"		
QY	959	CCACCAACCA					

RESULT 11

AE016766/C

LOCUS

DEFINITION

QY	359	AAAAAGACCGAAGCGGTGATCATCCACCTGGAACTGACACCACTGGAAGAGACCGCTT	418
Db	68968	GGGATAAACTGACGGCTTCGTCATTTACCCACGGTACCGACAGTGAAGAAACCGCTT	68909
QY	419	TTCTTCTCAACCTCACGGTGAAGCAAAACCTGCTGCGCTTTAGCGCCATCGCTC	478
Db	68908	ACTTCTTCACTGACGGTGAAGTGAAGCAAAACCTGCTGCGCTTTAGCGCCATCGCTC	68849
QY	479	CAGGCTTCTCCATGCTGCTGATGCGCCCATGATCTCTATAACCGCGTGAATGAGCGA	538
Db	68848	CGTCCACGCTCTAGCGCGCAGCGTCCATTTCAACCTGTATAACCGCGTGAATGAGCGA	68789
QY	539	TCAACAAAGCCTCTACTAAACAAAGGAGTGGTGTGATGATGAAGATGAGATTCAGCGC	598
Db	68788	CTGATAAGCCTCGCTTAATCGTGGGTACTGCTGATGATGAAGATGAGATTCAGCGC	68729
QY	599	CCGAGAGAGGCAACCAAGCTCAACACACCGCAGTCAATGCTTGTGCGCCCAACACAG	658
Db	68728	GCGGTGATGTACCAACAAACCAACACCGCAGTGTAGGACCTTCAAGTCTGTTAATACG	68669
QY	659	GTAAATCGGCACAGTCTATTATGCAAAAGTCGAGTATTCACTCAATCGGTTGACCTC	718
Db	68668	GTCTCTGGGTACTCATCAACACGTTAAGATTGACTACCAACGTTACCCCGCACGTAAGC	68609
QY	719	ACACCTTTCGAAGTGGATTGATATTAGCAAAATCGAAGACTCCCGAGAGTGCATATTC	778
Db	68608	ACACGAGGACACGCGCTTCGATGCTCTTAAGCTGAATGAATCTCCGAAAGTTCGCAATG	68549
QY	779	TTTACGCTACCCCGATGATGATGTTTGTAGTCAATGATGAGCGCTTCAGCGAGGACCA	838
Db	68548	TTTATACTACGCTAACGATCCGATCTTCGCGTAAAGCACTGATAGTGGCGCTATG	68489
QY	839	AAGGAATCATCATGACGCAATGGCAATGGCAACCTTTCCCTTTGACTCAAAATGCTC	898
Db	68488	ATGGCATGTTAGCGCTGGCGTGGTGAACGCAACCTGTATAAAACCGTATTTGACACCC	68429
QY	899	TTGAAAGAGGACCAATCAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	958
Db	68428	TTGCAACCGCTGCGAAACGCGCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	68369
QY	959	CCACACCAACGAGGCTGAAGTGGATGATGAAGAACTTTGTTTGGTGGTGGTGGTGGT	1018
Db	68368	CTACCACTCAGGATCGGAGTGGATGATGCGAAATAGCGCTTCGTCGCTCCGCGATGT	68309
QY	1019	TCACCTCAAAAGCAGAGTCTTATGTTAGTGGCTCACCACCAAACTAGTGATAGAG	1078
Db	68308	TGAACCGCAAAAGCGCGCTTCTGCTGCAACTGCTGCTGCAACTGCTGCTGCAACTG	68249
QY	1079	AGCGGATCCAAAGATCTTCTCCACCTATTATC	1112
Db	68248	AGCAGATCCAGCATCTTCAATCAGTACTAATC	68215
RESULT 12			
U32758			
LOCUS	10173 bp DNA linear BCT 29-MAY-1998		
DEFINITION	Haemophilus influenzae Rd section 73 of 163 of the complete genome.		
ACCESSION	U32758 L42023		
VERSION	U32758.1 GI:1573747		
KEYWORDS	Haemophilus influenzae Rd		
SOURCE	Haemophilus influenzae Rd		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.		
REFERENCE	1 (bases 1 to 10173)		
AUTHORS	Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G.G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Shrivley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,		
TITLE	Smith, H.O. and Venter, J.C.		
JOURNAL	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd		
MEDLINE	Science 269 (5223), 496-512 (1995)		
PUBMED	95350630		
REFERENCE	2 (bases 1 to 10173)		
AUTHORS	Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E. and Koonin, E.V.		
TITLE	Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli		
JOURNAL	Curr. Biol. 6 (3), 279-291 (1996)		
MEDLINE	96398784		
PUBMED	8805245		
REFERENCE	3 (bases 1 to 10173)		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	4 (bases 1 to 10173)		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes		
REFERENCE	5 (bases 1 to 10173)		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REMARK	The whole genome was shifted by 588 nucleotides for a new start		
COMMENT	On Sep 30, 1996 this sequence version replaced gi:1221457.		
FEATURES	Location/Qualifiers		
source	1. 10173		
gene	/organism="Haemophilus influenzae Rd"		
CDS	/mol_type="genomic DNA"		
	/db_xref="taxon:71421"		
	247..1704		
	/gene="HI0740"		
	247..1704		
	/gene="HI0740"		
	/note="similar to GB:M34393 SP:PI8159 PID:142994		
	PID:226139 GB:AL009126 percent identity: 38.72;		
	identified by sequence similarity; putative"		
	/codon_start=1		
	/transl_table=11		
	/product="phosphomannomutase (yhxB)"		
	/protein_id="AAC22400.1"		
	/db_xref="GI:1573748"		
	/translation="MGMRVLVSQAAGLAELKGYDKPSIVIGDCKNSDVFPARD		
	TAETMACGAVKYLPRKLPVLAIAIQFTDFTAGVMTVASHNPPDNQYKVLGKA		
	NGSQIVSPADKIDIALIDKVAIGNIOLPRSDNYVVLNDEVDAITTKTASLAKPA		
	LDINVTYMHGIVGYEVLKTLAKAGLPQPHVADQVMDPTFTVNFNPEEKGLD		
	LAIVAKENAEFIANDPADRLAVAPDAQGNKSLHGNVVCFLGVLAKYQYQK		
	QGTACSLVSPALAEIAKTYSPQSEETLIGFKYIGKVGILGFEFALGYLVDPKV		
	RDKDGI5AAIVFLDLVRLKKQKTLADYDEFTKEFGVAVSGQISIRVSDLSIKGL		
	MTALRNPPAEIAGVKAQFIDHIKTRQSDILVFNLENGRLTARFSGTEPKIFYL		
	DARGKDPKADRVLAEEFDEGVRHILRQDAYKQDC"		
	1853..2182		
	/gene="HI0742"		
	1853..2182		
	/gene="HI0742"		
	/note="hypothetical protein; identified by GeneMark;		
	putative"		
	/codon_start=1		
	/transl_table=11		
	/product="H. influenzae predicted coding region HI0742"		

```

/protein id="AAC22407.1"
/db xref="GI:1573749"
/translation="MCNLLVAEQCRICELENGWYNTYTESVPTLLANNAFYGSVEN
GKISSALRAELAAEAAAYAVALVQVGLGEGFAGLIAQMDMTSNGALTCLCKRWYLV
IKNLHLS"
complement (2249..2758)
/gene="HI0743"
CDS
/gene="HI0743"
/notes="similar to SP:P15040 GB:M24489 PID:147797
PID:466747 GB:U00096 percent identity: 56.16; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-export protein (secB)"
/protein id="AAC22401.1"
/db xref="GI:1573750"
/translation="MSEQKQDVAAETEEQOPVLIQIYRIYKVDVSFEAPNLFHQEWSK
PKLGFPLSTETQVGGDLYEVNLINISVETTLSDSDGVAFICEVKAQGVFTIGLEDDVQ
MAHLTSQCPNMLFVARELVSNLVRGTTPALNLSPVNFDALEFVEMRQQAENAE
KSEEEQTKH"
complement (2772..3218)
/gene="HI0744"
CDS
/gene="HI0744"
/notes="similar to SP:P37688 PID:466749 GB:U00096
PID:1790040 percent identity: 50.74; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein id="AAC22402.1"
/db xref="GI:1573751"
/translation="MOEFISMATPEAFQKHTLLAVSWFAIFVMVIVTFYKGTASKFKVI
THNEVRLINDEALIVDLRLSEEFQRGHIIINSVLPSIKNQNIKLESHKENALI
LVDTNGTSASASAVELTKQGENSVFLKEGLSAAVLAALPLVKKH"
3444..4493
/gene="HI0745"
CDS
3444..4493
/gene="HI0745"
/notes="similar to GB:M34234 SP:P00805 GB:M34277 PID:145277
PID:146597 percent identity: 70.52; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="L-asparaginase II (ansB)"
/protein id="AAC22403.1"
/db xref="GI:1573752"
/translation="MKLTKLALCTFLGLGVSIANAADLPNITILATGGTTAGSGQSSV
NSAYKAGQLSIDTLEIAPVPMKNIANIKGQIVKIGSQDMDNDEWLKLAIRACKS
TDGFVITHGTDTMEETAYFLDLTVKCBKPVLYGAMPATEKSDGPNLYNAVVA
DKSSGRGLVAMNNEVLGARDVTKTSTAVQTFHSPNYGSLGIHNSKVDYERSPE
KHTINTPNVEKLDLSLPKGLIIVAYSNAPVEPLNALINAGYOGIVSAGVGNVNAH
LDLEKAAKDSVVVRSSRFTGTTTRDAEVDSDSKYGFVASGTLPQKARVLLQALT
QTKDPKVIQGFDEF"
4649..5971
/gene="HI0746"
CDS
4649..5971
/gene="HI0746"
/notes="similar to GB:U14003 SP:P14409 GB:M27058 GB:X79886
PID:146047 percent identity: 42.59; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="anaerobic C4-dicarboxylate membrane transporter
protein (dcuB)"
/protein id="AAC22404.1"
/db xref="GI:1573753"
/translation="MSAMFLLQFAIVLCLIFVGARVGGIGLVFGGLGLIALISFGFL
KPAFLDFWFMIMAVAAAMAAAGGLDYMKIATILRRNPKHITFIAPAVTWLF
TLIAGTGHVAVSLVPIAVRQNGIRPERMSMAVIASOFAIVASPIAAVAVVAY
LEPGHILHGLDMVTIPSTILGLFLACLFPVKNWKEKLDKDDPEYORLNDPKADIPAS
TTSVEVEVSKTAKISVSLFLPGLLVLMGAAPSRIPVFDGKPMGMAHTIILMISI

```

```

GALLIITCKPDGTALTITKGSVFHAGMRAVIAIFGIAWLGDITLMQAHITETVKEMVKGLVE
TAPWAFAFALFVLSVNSQCAIVATLFPGLIAGLPAPVLIGVFPVAVNGFFPIYNG
PIIASLDFDTITGTTKIGKYILNHSFMLPGLLSMFFCLILIGLGLSHVIL"
6175..7509
/gene="HI0747"
CDS
6175..7509
/gene="HI0747"
/notes="similar to SP:P00393 PID:581140 GB:U00096
PID:1651546 PID:1787352 percent identity: 57.77;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase (ndh)"
/protein id="AAC22405.1"
/db xref="GI:1573754"
/translation="MKNVIVCGGAGGIELATFLGNLGRQKQAKVTLVDRNATHLWK
PILHIAITGMDGVDLSYRAHGKHHFSPQGSIRINREQKYVELAPVYQEGDM
LVIAARRIPYDLIVIAIGSKSNDFTKGVADNCIFLDSSKQALRFQHLKLEFLKFS
RALDDIGEEFKQKLVNDENKVNIAIVGGATGVLTAEYHAAEDLSSYGYKIDSSC
IQVTLVEAGTRLLPALPENLSAAVDELKEMGTNQLMTITEAOPTNLTITKGGBIK
ADLIWMAAGVRASVTVPQFDGLEINQLVVDLTQTTVDSDISFAIGDCAALIQSG
KLVPRAQAHHOMAKACAKNIFALPENPLKSPKNDKGLVSLNFTALGSLTNKFG
KNPLTVQSKFAGPAYVSLYRMHQHAGHCIKIGLILVLDKLNRYLKPRKLH"
complement (7589..10021)
/gene="HI0748"
CDS
complement (7589..10021)
/gene="HI0748"
/notes="similar to SP:P00482 GB:K00127 PID:147295
PID:148102 PID:148109 percent identity: 57.21; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glycerol-3-phosphate acyltransferase (plsB)"
/protein id="AAC22406.1"
/db xref="GI:1573755"
/translation="MANFINMYRQLLSLSALYKNNPIPANPIEELSNIHQPIVYV

```

```

Query Match      26.2%; Score 296.4; DB 1; Length 10173;
Best Local Similarity 57.1%; Pred. No. 1.1e-73;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

Qy 126 ACCCAAGTGAATCTCTAGCCAGGAGGACCATCGTGGTTCGGGGGAATCTAGCGT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3515 ACCAAATATTACAATCTTGGCAACGGGTGTACCATTCGAGGAGCGGCAAGTTCGGT 3574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 186 CAAGAGTAGTACTCTCTGAGCAGTCAACGGTTGATAAGCTTCTTCAGCCGCTCCCTGC 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3575 AAATTCGCGTATAAGCTGGCAATTAAGTATTGATACCTTAATTGAAGCTGTACCAGA 3634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 246 CATCAACGACTAGCCACCACCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3635 AATGAAAATATTGCCAACATTAAGGTGAGCAATTTGTAATAATAGTTTCACCAAGACAT 3694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 306 GACGGTAGAGTGTGGCTTAAACTAGCCAAAGCGTGTCAATGAGCTCTCTGCCCAAAAGA 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3695 GAATGACGAAGTCTGGCTAAACTGGCAAAAGCCATCAATGCTC-----AATGTAAG 3748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 366 GACCGAAGCGGTGATCATCACCCTGGAATGACACCATGGAAGAGACCGCTTCTTCCT 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3749 TACTGATGATTGTTCATTTACCATGTGACAGATACCATGGAAGAGACCGCTTCTTCCT 3808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 426 CAACCTCAGCGTGAAGAACCAAAACCTGTCTGCTCTGTAGCGGCGCATGCTCCAGGCTC 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3809 AGATTAAACCGTAAATGTGAAAACCGGTTCTCTGTTGGGCAATGGCTCTCAAC 3868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 486 TTCCATGAGTGTGATGCGCCCATGAATCTCTAAGCCCGTGAATGTAGCATCAACAA 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3869 AGAAAAAGTGTGATGCGCCCATTAATCTTTTACATGCTGCTGTTGTCGACGACACA 3928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 546 AGCTCTTACTAACAAAGGAGTGTGATTTGATGAACCATGAGATTACGCGCCGACAGA 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3929 AAAATCAAGTGTGCTGCTGTTTATTAGTCGAATGAATTAAGTACTAGGTGCTCGCGA 3988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	102782	AAATCTCGTATTAAGCTGGACAATTAAAGTATTGATCTTTAAATTGAAGCTGTACCAGA	102841
Qy	246	CAATCAACGACCTAGCCACCATCAAGGCTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT	305
Db	102842	AATGAAAAATATTGCGCAACATTAAGGAGGAGCAATTTGAAAAATAGGTTTCAAGACAT	102901
Qy	306	GACGGGTAGGTTGCTGCTTAACTAGCCAAAGCGTGTCAATAGCTCTCTCCGCCAAAAGA	365
Db	102902	GAATGACGAGTCTGGCTTAACTTGCACAAAGCCATCAATGCTC-----NATGTAAAG	102955
Qy	366	GACCGAAGCGGTGATCATCAACCATGGAATGACACATGGAAGAGACCGCTTTCTTCCT	425
Db	102956	TACTGATGATTTGTCATTTACCATGATACAGATACCATGGAAGAGACCGCTTTCTTCCT	103015
Qy	426	CAACCTCAGCGTGAAGAGCCAAAACCTGCTGCTGCTTGTAGGCGCCATGCTCCAGGCTC	485
Db	103016	AGATTTAAACCGTAAATGTGAAAACCGGTTGCTTCTGTTGGGCAATGCTCTGCAAC	103075
Qy	486	TTCCATCAGTCTGATGCGCCCAATGATCTCTATAAGCGCGTGAATGCTAGCGATCAACAA	545
Db	103076	AGAAAAAGTCTGATGCGCCCAATTAATCTTTACAATGCTGTGCTTGTGCGAGCAGACAA	103135
Qy	546	AGCCTCTACTAAACAAAGAGTGGTGAATGTGATGAACGATGAGATTCACGCCGCCAGAGA	605
Db	103136	AAAATCAAGTGGTGTGCTGTTTGTGCGCAATGAATAAAGTACTAGTGTCTCGCA	103195
Qy	606	AGCGACCAAGCTCAACACCCAGTCAGTCAATGATTTGCTTCGCCCAACACAGCTAAAT	665
Db	103196	TGTAACAAAACACAGTACGACCGCAGTCAACAGTTCATTCACCAAAATTAATGTTCTCT	103255
Qy	666	CGGCAAGCTTATTTATGCAAGTCTGATTTTCACTCAATCCGTTCCAGCTCAGCTC	725
Db	103256	AGGCTATATTCATACACGAAGTGGACTATGAACGTTCCCGAGAAACAAACATACCAT	103315
Qy	726	TGCAAGTGAATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATTTCTTTACGC	785
Db	103316	CAACACTCCGTTTAACTAGTAAATATTAGTACCTACCAAGTGGGATTTATTTATGC	103375
Qy	786	TCACCCCGATGATCTGATGTTTATGTCATGCGACCTTCAGGCGAGGAGCAAGGAAAT	845
Db	103376	TTATTCMAATGCACTGTGCAACCATTAACGATTAACGATTAACGATTAACGATTAACGAT	103435
Qy	846	CATCCATGAGGCGATGGCAATGGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGAATA	905
Db	103436	TGATCTGAGGAGTTGGCAATGGAAATGTTAATGCTGCACACTAGATCGCTTAGAAAA	103495
Qy	906	AGCAGCAAAATCAGGCGTGTGCTGCTGCAAGCTCTAGAGTGGGCGAGTGTTCACACAC	965
Db	103496	AGCGCAAAAGATAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	103555
Qy	966	CCAAGAGGCTGAAGTGAATGAAGAACTTTGGTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1025
Db	103556	TCGTGAGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	103615
Qy	1026	TCAAAAGCCAGAGTGTCTTATGTTAGCCCTCACCAAACTAGTGTAGATGAGAGCGGAT	1085
Db	103616	ACAGAAAGCCGCGTGTCTTTGCAATTAGCTTTTAACTCAAACTAAAGATCCNAAAGTAT	103675
Qy	1086	CCAAAAGATCTTCTCCACCTATTAA	1110
Db	103676	TCACCAATATTTGGAAGACTTCTAA	103700

RESULT 14	AR274513_08
Sequence split into 19 fragments	
LOCUS AR274513 Accession AR274513	
WPCOMMENT	
Fragment Name	Begin End
AR274513_00	1 110000
AR274513_01	100001 210000
AR274513_02	200001 310000
AR274513_03	300001 410000
AR274513_04	400001 510000

Qy	606	AGCGACCAAGCTCAACACCGCAGTCAATGCTTTGCTTGGCCCAACACAGGTAAAT	665
Db	3989	TGTAACAAAAACCGAGTACGACGCGAGTGAACAGCTTCATTTCAACCAATATTGTTCTCT	4048
Qy	666	CGGCAAGCTCTATTATGGCAAAAGTCTGATTTTCACTCAATCCGTTTCAGCTCAGCCCT	725
Db	4049	AGGCTATATTATACAGCAAGTGGACTATGAAGCTTCCCGCAGAAAGCAACATACAT	4108
Qy	726	TGCAAGTGAATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATTCTTTACGC	785
Db	4109	CAACACTCCGTTTAACTAGAGAAATTTAGATAGCTTACCCAAAGTGGGATTTATTATGC	4168
Qy	786	TCACCCCGATGATGATGATTTTGTAGTCAATGAGCCCTTTCAGGAGGAGCCAAAGGAT	845
Db	4169	TTATTCAAATGACCTGTGCAACCAATTAAGCAATTAATCAATGCTGCTGCTCAAGGAT	4228
Qy	846	CAATCAATGAGGATGGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTTGAAAA	905
Db	4229	TGATCTGAGGAGTTGGCAATGGAAATGTTAATGCTGCACTTTAGATCGCTTAGAAAA	4288
Qy	906	AGCAGCAATCAGGCGTGTGCTGCTGCAAGCTCTAGAGTGGGCGAGTGGTTCCACCAAC	965
Db	4289	AGCGCAAAAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4348
Qy	966	CCAAAGGCTGAATGATGATGAAGAACTTTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1025
Db	4349	TGCTGACGCTGAATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4408
Qy	1026	TCAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTGTAGAGAGGCGAT	1085
Db	4409	ACAGAAAGCCGCGTGTCTTGCATTTAGCTTTAACTCAAACTAAAGATCCNAAAGTAA	4468
Qy	1086	CCAAAAGATCTTCTCCACCTATTAA	1110
Db	4469	TCACCAATATTTGGAAGACTTCTAA	4493

RESULT 13
AR274513_07
WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1810121

Continuation (8 of 19) of AR274513 from base 700001 (AR274513 Sequence 1 from patent US

Query Match 26.2%; Score 296.4; DB 6; Length 110000;

Best Local Similarity 57.1%; Pred. No. 1.5e-73;

Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

Qy	126	ACCCCAAGTGAATCTTACCCAGAGGAGCAACATGCTGGTTCGGGGAATCTAGCT	185
Db	102722	ACCAATATTACATCTTTGCAACCGGTGTTACATTTGCAAGAGCGGCAAGTTCCGT	102781
Qy	186	CAAGAGTAGCTCTCTGAGGAGTCAACGCTTGATAGCTTTCTTGACGCGCTCCCTGC	245

AR274513_05 500001 610000
 AR274513_06 600001 710000
 AR274513_07 700001 810000
 AR274513_08 800001 910000
 AR274513_09 900001 1010000
 AR274513_10 1000001 1110000
 AR274513_11 1100001 1210000
 AR274513_12 1200001 1310000
 AR274513_13 1300001 1410000
 AR274513_14 1400001 1510000
 AR274513_15 1500001 1610000
 AR274513_16 1600001 1710000
 AR274513_17 1700001 1810000
 AR274513_18 1800001 1830121

Continuation (9 of 19) of AR274513 from base 800001 (AR274513 Sequence 1 from patent US

Query Match 26.2%; Score 296.4; DB 6; Length 110000;
 Best Local Similarity 57.1%; Pred. No. 1.5e-73;
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 126 ACCCAAGTACTATCTAGCCACGAGGAGGACCATCGCTGGTTTCGGGGGAATCTAGCGT 185
 DB 2722 ACCAAATATTACAAATCTTGCAACGGGTGTACCATTCGAGGAAGCGGCAAAATTCGGT 2781

QY 186 CAAGAGTACTCTCTGCTGGGACGCTACCGTTGTATAGCTTCTTTCGAGCGCTCCCTGC 245
 DB 2782 AATTTCTCGTATAAGCTGGGACATTAAGTATTGATCTTTAAATTGAAGCTGTACCA 2841

QY 246 CATCAAGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTCGCTCCCAAGAT 305
 DB 2842 AATGAAAAATATTGCCAATTAAGGTGAGCAATTTGAAAATAGGTTCAAGACAT 2901

QY 306 GACGGTAGGTGCTGCTTAACTAGCCAAAGCGTGTCAATGAGCTCTCGCCCAAAAGA 365
 DB 2902 GAATGACGAGTCTGGCTAAACTGGCAAAAGCCATCAATGCTC-----AATGAAAG 2955

QY 366 GACGAAGCGGTGATCATCCCATCGAAGTCAAGCATTCGAGGACGCTTCTTCTCT 425
 DB 2956 TACTGATGATTTGTCATTACCATTCGATGATGATGATGATGATGATGATGATGAT 3015

QY 426 CAACCTCAGGTGAAGAACCCAAACCTGTCTGCTTTAGCGGCGCATCGCTCCAGGCTC 485
 DB 3016 AGATTAAACCGTAAATGTGAAACCGGTTGTTCTGCTGGGCAATGCTCTGCAAC 3075

QY 486 TTCCATGAGTGTCTGATGCCCGCATGAATCTCTATACCGCTGAATGAGGATCAACAA 545
 DB 3076 AGAAAAAGTCTGATGCCCATTAATCTTTACATGCTCTGTTGTCGAGCAGACAA 3135

QY 546 AGCCTCTACTAAACAGAGTGGTGTATGATGAACGATGAGATTCAAGCTTCAAGCGCAG 605
 DB 3136 AAAATCAAGTGGTCTGCTGTTTGTAGTCGCAATGAATGAAGTACTAGTGTCTCGCA 3195

QY 606 AGCCACCAAGTCAACACCGCAGTCAATGCTTTGCTTCGCCCAACAGGATAAAT 665
 DB 3196 TGTAACAAACAGCAGTACGCGGAGTCAAGCGTTCATTCACCAATATGTTCTCT 3255

QY 666 CGGCAAGTCTATTATGGCAAGTTCGAGTATTTCACTCAATCGCTTCGACCTCACACCT 725
 DB 3256 AGGCTATTTCAACAGCAAGTGGACTATGAAGTTCCTCCGAAAGCAAAATACCAT 3315

QY 726 TGCAAGTGAAGTTGATATTAGCAAAATCGAAGTCTCCGAGAGTTCGATATTTCTTACGC 785
 DB 3316 CAACACTCCGTTTAAACGTAGAAAATTTAGATAGCTTACCCAAAGTGGGATTTATATGC 3375

QY 786 TCACCCGATGATCTATGATGTTTGTAGTCAATGAGCGCTTCAGGCGAGGAGCAAGGAT 845
 DB 3376 TTATTCAAATGCACTGTGCAACCAATTAACGCAATTAATGCTGGGTATCAAGGAT 3435

QY 846 CATCCATGAGGATGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGAATA 905
 DB 3436 TGATCTCGAGGATTTGGCAATGGAAATGTTAATGCTGCACACTTAGATCGCTTAGAAA 3495

QY 906 AGCAGCAAAATCAGGCGTAGTCTCGCTCGAAGCTCTAGAGTGGGAGTGGTTCACCAAC 965

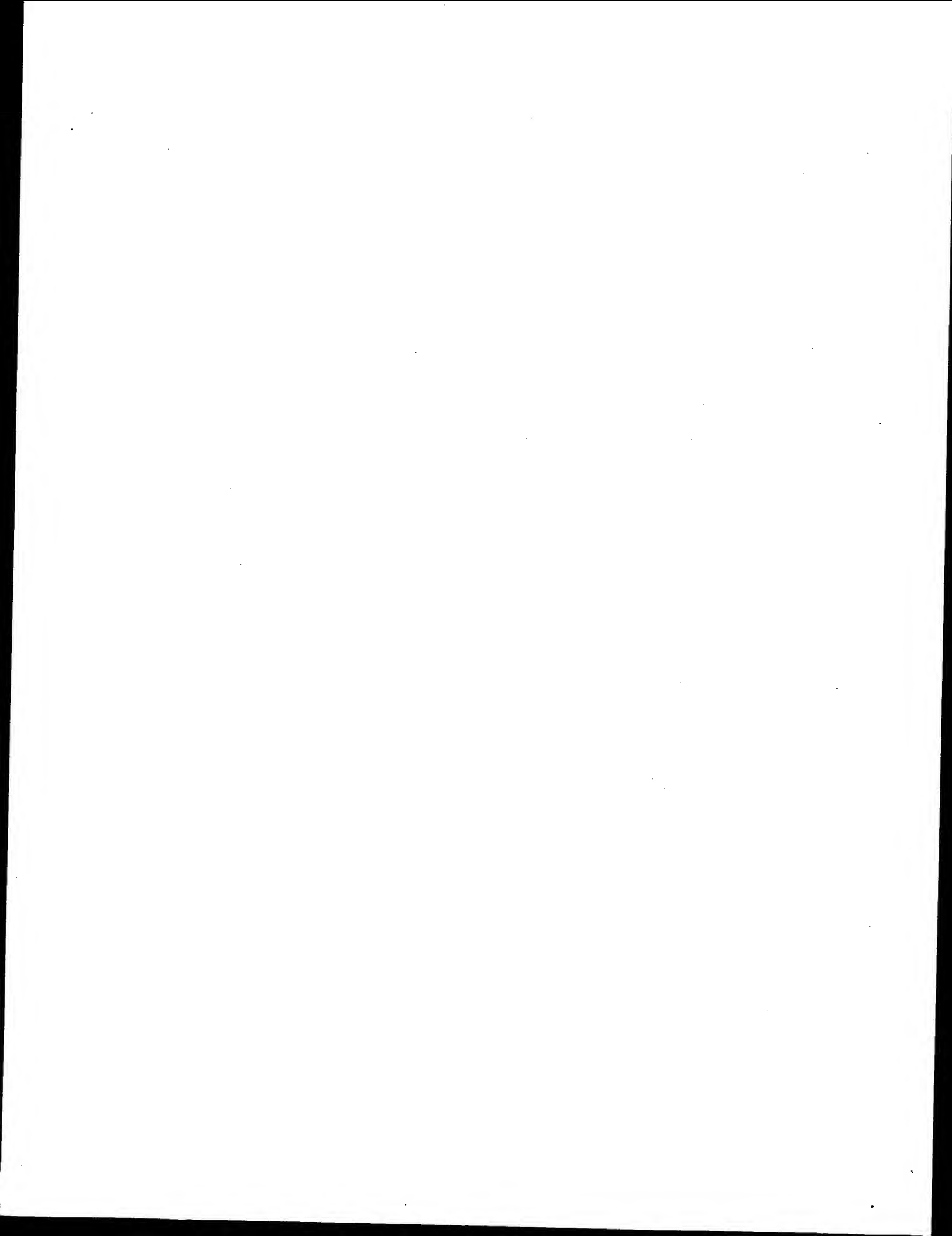
Db 3496 AGCGCAAGAGATAGCGTCTGTTAGTCGGTTCCTCTCGTGTACCAACGGGTATTACAA 3555
 QY 966 CCAAGAGCTGAAGTGGATGATAGAAACTTGGTTTGTGGCTACAGAGCTCTCAACCC 1025
 Db 3556 TCGTGACGCTGAAGTTCATGATAGTAATAATAGCTTTGTAGCATCAGGTACTTTAAATCC 3615
 QY 1026 TCAAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCAAACTAGTATGATGAGAGGCGAT 1085
 Db 3616 ACAGAAAGCCCGGCTCTCTGCAATAGCTTTAACTCAAACTAAAGATCCNAAAGTAAT 3675
 QY 1086 CCAAAAGATCTTCTCCACCTATTAA 1110
 Db 3676 TCAACAATATTTCGAAGACTTCTAA 3700

RESULT 15
 ECOLASNII
 LOCUS
 DEFINITION
 E.coli L-asparaginase II (ansB) gene, complete cds.
 ACCESSION
 M34277
 VERSION
 M34277.1 GI:146596
 L-asparaginase II.
 KEYWORDS
 Escherichia coli.
 SOURCE
 Escherichia coli
 ORGANISM
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 REFERENCE
 1 (bases 1 to 1530)
 JENNINGS, M.P. and BEACHAM, I.R.
 ANALYSIS of the Escherichia coli gene encoding L-asparaginase II,
 ansB, and its regulation by cyclic AMP receptor and FNR proteins
 J. Bacteriol. 172 (3), 1491-1498 (1990)
 90170867
 2407723
 COMMENT
 Original source text: E.coli (strain K-12) DNA.
 FEATURES
 Location/Qualifiers
 1..1530
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /db_xref="taxon:562"
 323..>1431
 /product="ansB mRNA"
 /evidence="experimental"
 385..1431
 /EC_number="3.5.1.1"
 /note="L-asparaginase II (ansB) precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AAA24062.1"
 /db_xref="GI:146597"
 /translation="MBFFPKTALAALVMPGFGAALALPNITILATGTTIAGGDSATK
 SNYTVGKGVENLVNAPQLKDIAVKGQVNVIGSDMNDVNLTLAKKINTDCDKT
 DGPVTHGDTMEETAYFLDLTKCDKPVVMGAMRPFSTMSADGPNLVNAYVTAD
 KASANGHVTVMNDTVDLDGRDVTKTNDVATFKSVNGPLGVIHNGKIDYQRTARK
 HTSDTPDKSLNELPKVGIYVYANASDLPAKALVDAGVDGIVSAGVNGNLYKSVF
 DLTATAAKTGTAIVRSSRVTGATTQDAEVDKYGCFVAGSGLNPKARVLLQALQTQ
 TKDPOIQIIFNQY"
 385..450
 /note="L-asparaginase II signal peptide"
 451..1428
 /product="L-asparaginase II"
 /EC_number="3.5.1.1"
 BASE COUNT 409 a 381 c 382 g 358 t
 ORIGIN

Query Match 26.1%; Score 295.2; DB 1; Length 1530;
 Best Local Similarity 56.8%; Pred. No. 1.7e-73;
 Matches 565; Conservative 0; Mismatches 423; Indels 6; Gaps 1;

QY 119 TGGCTAAACCCCAAGTACTATCTAGCAGGAGGACCATCGCTCGTTCGGGGGAT 178
 DB 446 TGGCAATTACCAATATACCATTTTAGCAACCGCGGAGCCATTCGCGTGGTGGTACT 505

QY	179	CTAGCGTCAAGAGTAGCTTACTCTCTCGAGCAGTCAACCGTTGATGAAGCTTCTTTGAGCGC	238
Db	506	CGGCAACCAATCTAACTACACAGTGGGTAAAGTTGGCGTAGAAAAATCTGGTTAATGCGG	565
QY	239	TCCCTGCCATCAACGACTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCC	298
Db	566	TGCGCAACTAAAGACATTGCGAACGTTAAGCGGAGCAGGTAGTGAATATCGGCTCC	625
QY	299	AAGAGATGACGGTAAAGTGTGGCTTAAACTAGCCAAAGCGGTCAATAGCTCTCTCGCC	358
Db	626	AGGACATGAACGATTAATGTCTGGCTGACACTGGCGGAAAAAATTAA-----CACGACT	679
QY	359	AAAAAGAGACCGAAGCCGTGATCATACCCATGGAATGACACCATGGAAGAGACCGCTT	418
Db	680	CGGATAAGACCGAGCGTCTGGTCATTATCCACCGGTACCGACACGATGGAAGAACTGC	739
QY	419	TCTTCTCAACCTCAGGTTGAAGGCAAAACCTGTCTGCTTTGTAGGCCCATCGCTC	478
Db	740	ACTTCTCGACCTGACCGGTGAATGCGCAAAACCGGTGGTGAATGGTTCGGCGCATGGTC	799
QY	479	CAGGCTCTTCCATGAGTGTGATGGCCCCCATGAATCTCTATAACCGCGTGAATGTACGA	538
Db	800	CGTCCACGTCATGACGCGACAGCGTTCATTCAACCTGTATTAACGCGTAGTACCGCAG	859
QY	539	TCAACAAAGCCTCTACTAACAAGAGTGGTGATTTGTATGAACGATGAGATTCACGCCG	598
Db	860	CTGATAAGGCTTCGCGCAACCGTGGGTCTGCTGTAGTGAATGACACCGTGTGTGATG	919
QY	599	CGAGAGAGCGACCAAGCTTCACACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAG	658
Db	920	GCGGTGAGTGCACCAAAACCAACACCGACGAGTACGACCTTCAGTCTGTTAACTACG	979
QY	659	GTAATAATCGCACAGTCTATTATGCGAAAGTCGAGTATTTCACTCAATCCGTTTCGACCTC	718
Db	980	GTCTCTTGGGTACATTCACAACGTTAAGATTGACTTACCAGGTATCCCGCGCACGTAAGC	1039
QY	719	ACACCTTTGCAAGTGAAGTTTGATATTAAGCAAAATCGAAGAACTCCCGAGTGCATATTC	778
Db	1040	ATACCAGCGACACGCCATTCCGATGTCTTAAGCTGAATGAATTCGCGAAAGTCGCGATTTG	1099
QY	779	TTTACGCTACCCCGATGATCTGATGTTTATGTCGAATGCAGCCCTTCAGGCAGGAGCCA	838
Db	1100	TTTATAACTACGCTTAACGCATCCGATCTTCGGCTTAAGACACTGGTAGATCGCGGTATG	1159
QY	839	AAGGAATCATCCATCGAGCATGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTC	898
Db	1160	ATGGCATCTGTAGCGCTGGTGGGTAAACGCAACCTGTATAAATCTGTGTTTCGACACG	1219
QY	899	TTGAAAAAGCGCAAAATCAGCGTAGTCTGCTCGCTCGAAGCTCTAGAGTGGGCAAGTGT	958
Db	1220	TGGCAGCCGCGGAAAAACCGGTACTGCGAGTGTGGCTTCTTCGCGGTACCGACGGCG	1279
QY	959	CCACCACCAAGAGAGCTCAAGTGAATGATAGAAACTTGGTTTTGTGGCTACAGAGATC	1018
Db	1280	CTACCACTCAGGATCGCGAGTGGATGATGCGAAATACGGCTTCGTCGCTCTCGCACGC	1339
QY	1019	TCAACCCCTCAAAAAGCCAGAGTCTTATGTTAGCCCTCACCAGAACTAGTGTATAGAG	1078
Db	1340	TGAAACCCCAAAAAGCGCGCTCTCTGCTGCAACTGGCTCTGACGCAAAACAAAGATCCGC	1399
QY	1079	AGCGGATCCAAAAGATCTTCTCCACCTAATATC	1112
Db	1400	AGCAGATCCAGAGATCTTCAATCAGTACTAATC	1433



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:23 ; Search time 441.315 Seconds

(without alignments)
6930.345 Million cell updates/sec.

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atggcagcagcatcatca.....aagaagggaattcttcttc 1133

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	20	AA03474 Wolinella succinog
2	1129.8	99.7	1133	21	AAC62512 W. succinogenes as
3	300.4	26.5	1044	21	AA032834 L-asparaginase DNA
4	296.4	26.2	1830121	17	AAT42063 Haemophilus influe
5	294	25.9	1848	18	AAT96346 Chimeric gene cont
6	259.2	22.9	1196	19	AAI4328 H. pylori GHPO 187
7	223.2	19.7	2837	8	AA070557 Sequence encoding
8	215.8	19.0	1014	15	AAQ68438 Pseudomonas glutam

ALIGNMENTS

RESULT 1

AA03474

ID AA03474 standard; DNA; 1133 BP.

AC AA03474;

XX 30-APR-1999 (first entry)

DE Wolinella succinogenes L-asparaginase DNA.

XX L-asparaginase; amplification; treatment; disease; asparagine depletion;
 KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW covalent modification; acylation; pharmacokinetic; immunogenic; spleen;
 KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;
 KW lymphocyte marker distribution; hepatotoxicity; ss.

XX Wolinella succinogenes.

XX WO9856410-A1.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US11905.

XX 09-JUN-1997; 97US-0049085.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

XX

Erwinia carotovora
 Helicobacter pylori
 CDNA encoding Tm 3
 CDNA encoding Tm 2
 CDNA encoding Tm 2
 CDNA encoding Tm 3
 CDNA encoding Tm 7
 pNOV4836 glucose i
 Salmonella sefa ge
 CDNA encoding Tm 2
 CDNA encoding Tm 3
 CDNA encoding Tm 7
 CDNA encoding Tm 2
 CDNA encoding Tm 1
 pNOV4835 glucose i
 CDNA encoding Tm 1
 Recombinant SOMgp5
 Insert from pRT8C
 S. pneumoniae type
 Streptococcus pneu
 Erwinia carotovora
 Genomic sequence o
 Streptococcus pneu
 Nucleotide sequenc
 S. epidermidis ope
 Staphylococcus epi
 S. epidermidis gen
 Erwinia chrysanth
 Nucleotide sequenc
 Enterococcus faeca
 Enterococcus faeca
 DNA encoding Hepat
 DNA encoding hepat
 Staphylococcus aur
 LFn-Bcl-Xl apoptos
 Nucleotide sequenc

CC	microbial enzymes asparaginase and glutaminease to treat autoimmune
CC	diseases and leukaemia. The diseases likely to respond to such treatment
CC	include graft versus host disease, rheumatoid arthritis, systemic lupus
CC	erythematosus, autoimmunity, collagen vascular diseases, AIDS,
CC	osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple
CC	sclerosis, sclerosing pancreatitis, rheumatic fever, inflammatory
CC	bowel diseases, primary biliary cirrhosis, chronic active hepatitis,
CC	glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves'
CC	disease. The present sequence is DNA insert used to demonstrate the
CC	method of the invention.
XX	Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 other;
XX	Query Match 99.7%; Score 1129.8; DB 21; Length 1133;
XX	Best Local Similarity 99.8%; Pred. No. 0;
XX	Matches 1131; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 ATGGCAGCAGCCATCATCATCATCATATAGCAGCGGCTGTGCGCGGCGAGCCAT 60
DB	1 ATGGGAGAGCCATCATCATCATCATATAGCAGCGGCTGTGCGCGGCGAGCCAT 60
QY	61 ATGGCTAGCATGCTGGTGGAGAGCAATGGTCCGGATCAGGGCTCTGTTTCATG 120
DB	61 ATGGCTAGCATGCTGGTGGAGAGCAATGGTCCGGATCAGGGCTCTGTTTCATG 120
QY	121 GCTAACCCCAAGTACTATCTAGCCAGGAGGACCATCGCTGGTTCGGGGGAATCT 180
DB	121 GCTAACCCCAAGTACTATCTAGCCAGGAGGACCATCGCTGGTTCGGGGGAATCT 180
QY	181 AGCGTCAAGTAGTACTCTCTGGAGCAGTACCGTTGATAGCTTCTGACCGGTC 240
DB	181 AGCGTCAAGTAGTACTCTCTGGAGCAGTACCGTTGATAGCTTCTGACCGGTC 240
QY	241 CTGCGCATCAAGCCTGACCCATCAAGGCTGAACAGATCTCAAGCATTTGGTCCCAA 300
DB	241 CTGCGCATCAAGCCTGACCCATCAAGGCTGAACAGATCTCAAGCATTTGGTCCCAA 300
QY	301 GAGATCAGGGTAAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCGCCCAA 360
DB	301 GAGATCAGGGTAAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCGCCCAA 360
QY	361 AAAGAGCCAGCGGTGATCATCACCATGGAAGTGAACAGATCAAGCATTTGGTCCCAA 420
DB	361 AAAGAGCCAGCGGTGATCATCACCATGGAAGTGAACAGATCAAGCATTTGGTCCCAA 420
QY	421 TTCTCTCAACCTCAGGTGAAAGCCAAAACCTGTGTGCTGTAGCGCCATGCGTCA 480
DB	421 TTCTCTCAACCTCAGGTGAAAGCCAAAACCTGTGTGCTGTAGCGCCATGCGTCA 480
QY	481 GGCTCTTCCATGAGTGTGATGCGCCCATCAATCTCTATACCGCTGATGTAGCGATC 540
DB	481 GGCTCTTCCATGAGTGTGATGCGCCCATCAATCTCTATACCGCTGATGTAGCGATC 540
QY	541 AACAAAGCCTCTACTAAACAAAGGAGTGGTGAATGTGATGAACGATGATTCAGCGGCC 600
DB	541 AACAAAGCCTCTACTAAACAAAGGAGTGGTGAATGTGATGAACGATGATTCAGCGGCC 600
QY	601 AGAGAGGAGCAAGCTCAACACACCGAGTCAATGCTTCTGCGCCCAACACAGGT 660
DB	601 AGAGAGGAGCAAGCTCAACACACCGAGTCAATGCTTCTGCGCCCAACACAGGT 660
QY	661 AAAATCGGCACAGTCTATTATGCAAAAGTCAGTATTTCACTCAATCCGTTCCAGCTCAC 720
DB	661 AAAATCGGCACAGTCTATTATGCAAAAGTCAGTATTTCACTCAATCCGTTCCAGCTCAC 720
QY	721 ACCCTTGAAGTGGATTTAGTCAAAATCGAAGAACTCCCGCAGTCTGATATTTCTT 780
DB	721 ACCCTTGAAGTGGATTTAGTCAAAATCGAAGAACTCCCGCAGTCTGATATTTCTT 780
QY	781 TAGCGTCAACCCCGATGATCTGATCTTTTAGTCAATGAGCGCCCTTCAGGCGAGGCCAAA 840
DB	781 TAGCGTCAACCCCGATGATCTGATCTTTTAGTCAATGAGCGCCCTTCAGGCGAGGCCAAA 840

Db 62 TGGCATTACCAATATCACCATTGTTAGCAACCGCGGACCATTCGCGTGGTGGTACT 121
 Qy 179 CTAGCGTCAAGAGTAGTCTCTCTGAGAGAGTCAACCGTGGTGAAGTCTTTCGACCG 238
 Db 122 CCGCAACCAATCTAATCTACACAGCGGTAAAGTTGGCTAGAAAATCTGTTAATCGG 181
 Qy 239 TCCTGTCATCAACGAGTCAAGCAGTCAAGCAGTCAAGCAGTCAAGCAGTCAAGCAGT 298
 Db 182 TGCGCAACTAAAGACATTCGCAACCTTAAAGCGAGAGGAGTAGTAATATCGGTCC 241
 Qy 299 AAGAGATGACGGTAAAGTGTGGCTTAAATAGCAAGCGTGTCAATGAGCTCTCGCCC 358
 Db 242 AGCAGATGAACGATAATGTCTGCTGACACTGCGCAAAAAATTA-----CACGACT 295
 Qy 359 AAAAAAGAGACCGAAGCGGTGATCATCAACCCATGGAATGACACCATGGAAGAGACGCTT 418
 Db 296 GCGATAAAACCGAGCGCTTCGTCTATACCCACCGGTACCGACAGATGGAAGAACCGCTT 355
 Qy 419 TCTTCTCAACCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 478
 Db 356 ACTTCTGACCTGACCGTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 415
 Qy 479 CAGGCTCTTCCATGAGTGTGAGTGGCCCATGATCTTATACCGCGTGAATGTAGCA 538
 Db 416 CGTCCAGTCCATGAGCGGAGAGCGGTCCATCAACCTGTATACCGCGTGTAGCCGAG 475
 Qy 539 TCACAAAGCTCTACTAAACAAAGAGTGTGATGTTGATGAACGATGAGATTCACGCG 598
 Db 476 CTGATAAAGCTTCGCTAATCTGCGGTGCTGTTGATGATGAACGATGAGATTCACGCG 535
 Qy 599 CCAGAGAGCGACCAAGCTCAACACACCGGAGTCAATGCTTCTGCGGTGATGTTGCGGCAACAG 658
 Db 536 GTCGCGATGTACCAAAACCAACACACCGGAGTCAATGCTTCAAGCTGTGTTACTAG 595
 Qy 659 GTAAATCGGCACAGTCTATTATGCAAGTGTGATGTTTCACTCAATCCGTTGACCTC 718
 Db 596 GTCTCTGGGTATACATTACCAACGTTAGATGATGATGATGATGATGATGATGATGATG 655
 Qy 719 ACACCTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
 Db 656 ACACAGGATACGCGATTCGATGCTCTAAGCTGAATGAGTGTGCGAAGTTCGGCATC 715
 Qy 779 TTTACGCTCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
 Db 716 TTTATACTAGCTAACGCTCCATCTTCCGGCTAAAGCACTGTGATGTCGGGCTATG 775
 Qy 839 AAGCAATCATCGAGGATGAGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTC 898
 Db 776 ATGGCATCGTTAGCGCTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATG 835
 Qy 899 TTGAAAAGCAGCAAAATCAGGCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
 Db 836 TGGCAACCGCGGAAAACGGCACTGAGTAGTGGCTTTCTTCCCGGTACCGACGGGTG 895
 Qy 959 CACACCCCAAGAGGCTGAAGTGAATGAAGAACTTGGTTTGTGGCTTACAGAGATC 1018
 Db 896 CTACCACTACGATGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATG 955
 Qy 1019 TCAACCTCAAAAGCAGAGTCTTCTTATGTTAGCTTACGCTTACGCTTCTGCTGCTT 1078
 Db 956 TGAACCGGAAAAGCGCGGTCTGCTGAGTGGCTCTGAGCGCAAAACCAAGATCCCG 1015
 Qy 1079 AGCGATCCAAAGATCTTC 1098
 Db 1016 AGCAGATCCAGCAGATCTTC 1035

RESULT 4
 AAT42063
 ID AAT42063 standard; DNA; 1830121 BP.
 XX
 AC AAT42063;

XX 14-SEP-1999 (first entry)
 XX Haemophilus influenzae complete genome sequence.
 XX Genome; bacterium; Haemophilus influenzae; computer readable medium;
 XX expression modulating fragment; regulation; gene expression; vector;
 XX organism; open reading frame; ORF; ds.
 XX Haemophilus influenzae.
 XX WO9633276-A1.
 XX 24-OCT-1996.
 XX 22-APR-1996; 96WO-US05320.
 XX 07-JUN-1995; 95US-0487429.
 XX 21-APR-1995; 95US-0426787.
 XX 07-JUN-1995; 95US-0476102.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
 XX WPI; 1996-485782/48.
 XX Haemophilus influenzae Rd genome recorded on computer readable
 XX medium - useful for identifying commercially important nucleic acid
 XX fragments by homology searching
 XX Claim 1; Page 77.2-77.1091; 1291pp; English.
 XX This sequence represents the complete genome sequence of the bacterium
 XX Haemophilus influenzae strain Rd. The invention relates to a computer
 XX readable medium (CRM) having recorded upon it the complete H. influenzae
 XX nucleotide sequence (I), a representative fragment of (I) or a nucleotide
 XX sequence at least 99% identical to (I). By providing the full-length
 XX genomic sequence in a computer readable form, it is possible to identify
 XX commercially important nucleic acid fragments and expression modulating
 XX fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
 XX regulate the expression of a nucleic acid molecule. Vectors and altered
 XX organisms comprising the predicted ORFs can be used to produce any of the
 XX polypeptide fragments of the H. influenzae Rd genome.
 XX SQ Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
 Query Match 26.2%; Score 296.4; DB 17; Length 1830121;
 Best Local Similarity 57.1%; Pred. No. 8.1e-81;
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
 Qy 126 ACCCAAGTGACTATCTCTAGCCACAGAGGACCATCGTGTTCGGGGAATCTAGCGT 185
 Db 802722 ACCAAATATTACAAATCTTGCAACGGGTGGTACATTGACGGAAGCGGCAAGTTCGGT 802781
 Qy 186 CAAGAGTAGTACTCTCTGCTGGAGCAGTCCCGTTGATAGCTTCTTGCAGCCCTCCCTGC 245
 Db 802782 AAATTCGCTATAAGCTGGACATTAAGTATTGATCTTAATTAAGAGCTGTACAGA 802841
 Qy 246 CATCAACGACCTAGCCACCATCAAGGTTGAAAGATCTCAAGCATTTGGTCCCAAGAGAT 305
 Db 802842 AATGAAAAATATTGCAACATTAAAGGTGAGCAAAATTTAAATAGGTTCAACAGCAT 802901
 Qy 306 GACGGTAAAGTGTGGCTTAAACTAGCAAGCGTGTCAATGAGCTCTCGCCCAAAAAGA 365
 Db 802902 GAATGACGAGTCTGGCTAAACTGCAAAAGCCATCAATGCTC-----AATGTAAGA 802955
 Qy 366 GACCGAAGCGGTGATCATCACCATTGGAATCTGACACCATGGAAGAGACCGCTTTCTTCT 425
 Db 802956 TACTGATGATTTGTCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 803015
 Qy 426 CAACCTCACGGTGAAAGACCAAAACCTCTCTGCTTGTAGCGCCATGCTCCAGGCTC 485

[illegible]

This sequence represents a polynucleotide of the invention. It was isolated from *Helicobacter pylori* and encodes a *H. pylori* GHPo protein. The polypeptide can be used for preventing or treating *Helicobacter* infections, and gastrointestinal diseases associated with these infections, including chronic, and atrophic gastritis, and peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used for the production of antibodies. The products can also be used for detection and diagnosis.

Sequence 1196 BP; 355 A; 222 C; 286 G; 333 T; 0 other;

Query Match	22.9%;	Score 259.2;	DB 19;	Length 1196;
Best Local Similarity	55.5%;	Pred. No. 7,3e-71;		
Matches 548;	Conservative 0;	Mismatches 428;	Indels 12;	Gaps 2;
Qy	126	ACCCCAAGTCACTATCTTAGCCACAGGAGGCACCATCGCTGTTCGGGGGAATCTAGCGT	185	
Db	157	ACCCACCATGTGTTTACTGGCCACAGGGGGACGATTTCAGGGAGTGTGCGAGGGCGAG	216	
Qy	186	CAAGAGTACTACTCTGCTGGAGCAGTCAACGTTGATAGCTTCTTGACGCGTCCCTGC	245	
Db	217	TTTGGGTAGTTATAGAGTGTGTAGTTGGGCATCAAGAGGCTTTTGAAGGCTATCCCTAG	276	
Qy	246	CATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCAATTCGCTCCCAAGAGAT	305	
Db	277	TCCTTAACAGACTCGCTCGCATTCAGGGGAGCAGTTTCTAACATCGGCTCACAAGACAT	336	
Qy	306	CACGGGTAAAGTGTGGCTTAACTAGCCACAGCGTGTCAATAGAGCTCTCGCCCAAAAGA	365	
Db	337	GAATGAAGAGGTATGGTTCAAGCTCGGCCAAACGTCGCCAAGAAATTGCTAGATGATAGCCG	396	
Qy	366	GACCGAGCGGTGATCATCAATCCCATGGAATCTGACACCATGGAAGAGACCGCTTCTTCCT	425	
Db	397	TATTCGAAGGCGTGTCTATCAGCATGSCACGGACATTTAGAAAGAGCGGTATTTTTT	456	
Qy	426	CAACCTCACGGTGAAAAGCCAAAACCTGTCTGCCTTGTAGGGCCATCGTCCAGGCTC	485	
Db	457	AAACTTAGTTTACGCTCCAAAAACCGTCTGCTGGTGGAGCGATCGTAAATGCTGC	516	
Qy	486	TTCCATGAGTGTGATGGGCCCATGAATCTCTAATACGCGGTGAATGTAGCGATCAACAA	545	
Db	517	TTCTTTGAGCGCGGATGGGCGTTTGAATTTATATATGTGTGACGGTACGCTCAATGA	576	
Qy	546	AGCCTCTACTTAAACAAGGAGTGGTGAATTGTGATGAACGATGAGATTCACGCCACAGAGA	605	
Db	577	AAAAGTGCATTAAGGCGGTGTATGTGTGTATGACGATAATATTTTATGCGCTAGAGA	636	
Qy	606	AGCGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAAT	665	
Db	637	AGTGAATTAAAAACGACACACCCACATTCACCTTTAAAGCCTTTAAATGCGGCGGAT	696	
Qy	666	CGGCAAGTCTAATATGCAAAAGTCGATATTTCACTCAATCCGTTGACCTCACACCT	725	
Db	697	AGGAGCGGTATATATGCAAAACCGGCTATTACATGCAAGCTTTTGAAAAACACACAC	756	
Qy	726	TGCAAGTGAAGTTTGATATATAGCAAAATCGAA--GAACTCCCCAGATCGATATCTTTTA	782	
Db	757	AGAGAGCGAAATTTTCCCTTTCAACACTCAAAACCCCCCTGCTTAAGTGGATATTATTA	816	
Qy	783	CGCTCACCCGATGATCTGATGTTTTTAGTCAATGACGCGCTTCAGGACGAGCCAAAGG	842	
Db	817	CACGATGCTGGCATGACCCCTGATTTTATCCAAGCGAGCCTAACTCCGATCAAAAGG	876	
Qy	843	AATCATCATCGAGGCATGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGA	902	
Db	877	CGTTGTGATACCGGGTGGTAATGGGAATGTGACGTGGGTTTTTAAAGCGATGCA	936	

P-PSDB; AAW98609;

Db
877. CGTTGTGATAGCCGGGTGGTGAATGGGAATGTGAGCGCTGGGTTTTTAAAGCGATGCA 936

QY	903	AAAGCAGCAAAATCAGCGCTAGTCGCTCGAAGCTCTAGAGTGGCGAGTGTCCAC	962
Db	937	AGAGCGAGCAAAATGGGGTGGTTATGTTCTTACGAGGTAATAATAGCGGTGAGAT	996
QY	963	CACCAAGAGCGCTGAAGTGATGATAAGAACTTTGGTTTGTGGCTACAGAGTCTCAA	1022
Db	997	TA-----CTTCAGCGCAGATTGATGACAGGCGCTTCATCAAGCGCACAATTTAAA	1047
QY	1023	CCCTCAAAAGCCAGAGTGTCTTATGTTAGCCCTCACAAAACCTAGTGTATAGAGGCG	1082
Db	1048	CCCCCAAAAGCTAGGGTGTCTTTTACAACTGCTTTAACTAAACAATAATAAGAAA	1107
QY	1083	GATCCAAAAGATCTTCTCCACCTATTAA	1110
Db	1108	AATCCAGAATGTTGAAGAGTATTGA	1135
XX	RESULT 7		
ID	AA70557		
XX	AA70557 standard; DNA; 2837 BP.		
AC	AA70557;		
XX	25-MAR-2003 (updated)		
DT	07-AUG-1991 (first entry)		
XX	Sequence encoding a protein with L-asparaginase (LA) activity in		
DE	PASN 30 or PASN 32.		
XX	Enzyme; cancer therapy; leukaemia therapy; ss.		
OS	Erwinia chrysanthemi.		
XX	Key Location/Qualifiers		
FH	misc_feature 1..25		
FT	/tag= a		
FT	/label= pUC9 linker		
FT	661..723		
FT	/tag= b		
FT	724..1707		
FT	/tag= c		
FT	/note= "claimed"		
FT	1772..1789		
FT	/tag= d		
XX	EP211639-A.		
PN	25-FEB-1987.		
XX	04-AUG-1986; 86BP-0305984.		
PF	06-AUG-1985; 85GB-0019753.		
PR	(PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.		
XX	Ackinson A, Minton NP, Gilbert HJ;		
XX	WPI; 1987-051784/08.		
XX	P-PSDB; AAP70348.		
DR	New recombinant plasmids coding for L-asparaginase - esp. from		
PT	Erwinia carotovora, useful for treating malignancies, and new		
PT	transformed hosts		
XX	Disclosure; Fig 4; 35pp; English.		
XX	When E.carotovora SCRI 193 contg. plasmid PASN 32 was cultured, the		
CC	cell-free culture broth contained 11.19 LA units/ml with specific		
CC	activity 53.34 units/mg of protein. These figures compare with 10.87		
CC	LA units/ml and 14.97 units/mg of protein for the currently used		
CC	prodn. strain E.chrysanthemi NCPPB 1066.		
CC	(Updated on 25-MAR-2003 to correct PA field.)		

XX	Sequence	2837 BP; 627 A; 665 C; 867 G; 678 T; 0 other;	
SQ	Query Match	19.7%; Score 223.2; DB 8; Length 2837;	
	Best Local Similarity	53.3%; Pred. No. 2.7e-59;	
	Matches	527; Conservative 0; Mismatches 443; Indels 18; Gaps 2;	
QY	127	CCCCAAGTGAATCTCTAGCCACAGGAGGACCATCGCTGTTCGGGGGAATC---TAGC	183
Db	736	CCCAATATCTTTCTTGGGACCGGGGTACAAATTCGGCTCAGCGCAACGGTACC	795
QY	184	GTCAGAAGTAGTACTCTCTGTGGAGAGTACACCGTTGATTAAGTCTTTGACGCGTCCCT	243
Db	796	CAAAACACAGGTTACAAAGCTTGGCGCTTGGCGTGGATACGCTTAATCAACGCTGTGCT	855
QY	244	GCCATCAACGACCTAGCACCATCAAGGGTGAACAGATCTCAAGCATTTGGTCCCAAGAG	303
Db	856	GAGGTGAAGAACTGGCTTAATGTGAAGGGGAGCAGTTCTCCAACATGGCCAGGAAAC	915
QY	304	ATGACGGGTAAAGTGTGGCTTTAAACTAGCCAAAGCGTGTCAATAGCTCTCTGCCCAAAA	363
Db	916	ATGACCGGTGATGTGGTCTCAAGCTAGCCAGCGTGTGAATGAATCTCTGCACCGGAT	975
QY	364	GAGACCGAAGCGGTGATCATCAACCATGGAATGACACCATGGAAGAGACCGCTTCTTC	423
Db	976	GATGTGATGTGTGGTATCACCCACGGGACCGACACGGTGGAAAGATCGGCTTACTTT	1035
QY	424	CTCAACCTCAGGTGAAGGCAAAACCTCTCTGCTTTAGGCGCCATCGTCCAGGC	483
Db	1036	CTTCACTGACGGTAAAGATGACACGAGCTGGTGTGTTGTCGACGATGCTGTCGGCA	1095
QY	484	TCTTCATGAGTGTGATGGCCCATGAATCTCTATAACCGCTGAATGTAGCGATCAAC	543
Db	1096	ACGGCCATCAGTGTGACGGCCGATGAACCTCTGGAAGCGGTACGCGTGGGGTGAC	1155
QY	544	AAAGCTCTACTAAACAAAGAGTGTGATGTGATGAACGATGATTCACGCGCGCAGA	603
Db	1156	AAACAGTCTCGCGTTCGGCGGTGATGGTGGTCTTATGATCGTATCGCTCTGCCGC	1215
QY	604	GAAGCGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGTAA	663
Db	1216	TACATCACCAAGACCAACGCTCTAGCTGGATACGTTCAAGCGCAATGAAGAAGGTAC	1275
QY	664	ATCGGCACAGTCTTATATGGCAAGTCTGAGTATTTCACTCAATCCGTTGACCTCACCC	723
Db	1276	CTGGCGTCAATTTGGTAAACCGCATTTACTACCAAAACCGTATCCAGCTGCATACC	1335
QY	724	CTTGCAAGTGAGTTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATTTCTTAC	783
Db	1336	ACCGGTCTGTTCGACGTGGTGGCTGACCTTCGCTCCGAAAGTCGACATTTCTTAT	1395
QY	784	GCTCACCCCGATGATCTGATGTTTATGTCATATGACGCCCTTCAGGAGGAGCCAAAGGA	843
Db	1396	GGCTATCAGGATGACCCGGAATATCTGTATGACGCGCTATCCAGCATGGCGTAAAGGT	1455
QY	844	ATCATCCATGCGAGCATGGCAATGGAAACCTTCCTTTGACTCAAAATGCTCTTGAA	903
Db	1456	ATCGTCTATGCGGTATGGCGCAGGTTCAGTCTCGCTTCGCGTATTCGCGGTATGGC	1515
QY	904	AAAGCAGCCAAATCAGCGGTAGTCGCTGCTGCAAGCTCTAGAGTGGCGAGTGGTCCACC	963
Db	1516	AAGCGATGGAGAAAGCGGTTGTGTGATCCGTTCTACCGCACAGGCAATGGTAT----	1571
QY	964	ACCAAGAGGCTGAAGTGGATGATAGAAACTTGGTTTGTGGCTACAGAGAGTCTCAAC	1023
Db	1572	-----TGTCGCGCGGATGAAGAGCTGCCAGGTCTGGTTCTTGACTCTCTTAAC	1620
QY	1024	CCTCAAAAGCCGAGGTCTTCTTATGTTAGCCCTCACAAAACCTAGTGTATAGAGGCG	1083
Db	1621	CCGCAATGCGCGCATCTTGTGATGCTGCAATGACTCGCACCAAGTATCCGAAAGTCC	1680
QY	1084	ATCCAAAAGATCTTCTCCACCTATTAAAT	1111

1

1

CC Helicobacter pylori selected from 421 complexes given in the
 CC specification. The complex of protein-protein interactions are useful
 CC for screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence encodes a selected interacting domain (SID), identified via
 CC protein-protein interactions.
 XX
 SQ Sequence 345 BP; 90 A; 69 C; 104 G; 82 T; 0 other;

Query Match 9.6%; Score 108.8; DB 24; Length 345;
 Best Local Similarity 57.7%; Pred. No. 1.1e-23;
 Matches 194; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 142 CTAGCCACAGGAGGACCATCGCTGGTTCGGGGATCTAGCGTCAAGAGTAGCTACTCT 201
 Db 2 CTGGCGACAGGGGACGATTCAGGGAGTGTGCGAGCGCGAGTTGGGTAGTTATAAG 61

QY 202 GCTGGAGGAGTCACCGTTGATTAAGCTTCTTGAGCGCGTCCCTGCCATCAACGACCTAGCC 261
 Db 62 AGTGTGTAGTTGGGCATCAAGAGCTTTTGAAGGCTATCCCTAGTCTTAAACAGACTCGCT 121

QY 262 ACCATCAAGGTTGAACAGATCTCAAGCATTCGCTCCCAAGAGATGACGGTAAAGTGTGG 321
 Db 122 CGCATTCAGGGGACGAGTTTCAACATCGGCTCACAAGACATGAATGAAGAGGTATGG 181

QY 322 CTTAAACTAGCAAGCGTGTCAATGAGCTCTCGGCCCAAAAGAGACCGGCGTATC 381
 Db 182 TTCAAGCTCGCCAAAGTGCACCAAGATTGCTAGATGATAGCGGTATTCAGGCGTGTGTC 241

QY 382 ATCAACCATGAAGTACACATGACAGACCGCTTCTTCCCTCAACCTCAGGTGAAA 441
 Db 242 ATCAAGCATGCGACGACACATTTAGAGAGAGCGGTATTTTAAACTTAGTTTATAGC 301

QY 442 AGCCAAAACCTGTCTGCTTGTAGCGCCCATGCGT 477
 Db 302 TCCAAAACCGTCTGCTGTGGAGCGATGCGT 337

RESULT 11
 AAS21124
 ID AAS21124 standard; cDNA; 543 BP.
 AC AAS21124;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding Tm 3.9/His tag fusion protein minus signal peptide.
 XX
 KW Anti-freeze peptide; Tm 3.9; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatisation; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.
 XX
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..513
 FT /*tag= a
 FT /product= "Tm 3.9 clone protein"
 FT misc_feature 64..165
 FT /*tag= b
 FT /note= "Nucleotides encoding His tag"
 FT mat_peptide 166..510
 FT /*tag= c
 XX
 WO200194378-A1.
 XX
 PN 13-DEC-2001.
 XX
 PD 07-JUN-2001; 2001WO-US18532.
 XX
 PF 08-JUN-2000; 2000US-210446P.
 XX
 PR

XX
 PA (UNYV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORNWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;
 XX
 DR WPI; 2002-090137/12.
 DR P-PSDB; AAU10053.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 XX
 PS Claim 4; Page 345; 364pp; English.
 XX
 CC This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protein to limit tumour cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatisation. The Tm 12.86
 CC antibody/antiserum which is also used as a screening device to screen
 CC cDNA libraries in an expression system, including cross-species cDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents a cDNA encoding the Tm 3.9 clone/His
 CC tag fusion protein minus the signal peptide of the invention. This cDNA
 CC was created to facilitate purification and to try to enhance the anti-
 CC freeze activity of the recombinant protein.
 XX
 SQ Sequence 543 BP; 164 A; 133 C; 144 G; 102 T; 0 other;

Query Match 9.1%; Score 103.6; DB 24; Length 543;
 Best Local Similarity 88.9%; Pred. No. 5.9e-22;
 Matches 112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGGACAGCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
 Db 64 ATGGGACAGCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 123

QY 61 ATGGCTAGCATGACTGGTGGACAGCAATGGGTGGCGGATCCAGCGCCCTCTGTTTCATG 120
 Db 124 ATGGCTAGCATGACTGGTGGACAGCAATGGGTGGCGGATCCCTCACCAGTGAACAGATA 183

QY 121 GCTAAA 126
 Db 184 CAGAAA 189

RESULT 12
 AAS21116
 ID AAS21116 standard; cDNA; 543 BP.
 XX
 AC AAS21116;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding Tm 2.2 clone/His tag fusion protein minus signal sequence.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:41:09 ; Search time 105,214 Seconds
(without alignments)
4753,042 Million cell updates/sec

Title: US-09-937-982-3
Perfect score: 1133
Sequence: 1 atggcagcagccatcatca.....aagaaagggaattctttccac 1133

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfileseq1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	3	US-09-094-435-3
2	296.4	26.2	1830121	4	US-09-557-884-1
3	296.4	26.2	1830121	4	US-09-643-990A-1
4	294	25.9	1848	1	US-08-447-422-15
5	215.8	19.0	1014	5	PCT-US92-10421-1
6	210.8	18.6	1017	4	US-08-050-482A-1
7	200.4	17.7	1182	4	US-09-252-991A-5793
8	200.4	17.7	11335	4	US-09-252-991A-5877
9	159.2	14.1	1125	4	US-09-328-352-2336
10	147.4	13.0	2748	4	US-09-252-991A-5773
11	101.4	8.9	552	4	US-09-230-078A-3
12	80	7.1	270	4	US-09-252-991A-5837
13	76.6	6.8	1005	4	US-09-107-532A-1293
14	74	6.5	2231	4	US-08-961-527-265
15	72.6	6.4	366	4	US-09-613-303-11
16	71.4	6.3	981	4	US-09-134-001C-397
17	70.4	6.2	1230	4	US-09-613-303-54
18	66.2	5.8	2847	4	US-09-613-303-20
19	65.2	5.8	1989	4	US-09-207-388-4
20	64.4	5.7	2016	3	US-09-198-723A-110
21	64.4	5.7	2016	3	US-09-198-723A-111
22	63.8	5.6	1624	4	US-09-647-224A-13
23	63.8	5.6	1707	4	US-09-207-388-10
24	62.8	5.5	648	3	US-09-198-723A-112
25	62.8	5.5	648	3	US-09-198-723A-113
26	62.8	5.5	648	3	US-09-198-723A-114
27	62.8	5.5	648	3	US-09-198-723A-115

Sequence 92, Appl
Sequence 93, Appl
Sequence 94, Appl
Sequence 95, Appl
Sequence 96, Appl
Sequence 97, Appl
Sequence 98, Appl
Sequence 99, Appl
Sequence 100, Appl
Sequence 101, Appl
Sequence 102, Appl
Sequence 103, Appl
Sequence 104, Appl
Sequence 105, Appl
Sequence 106, Appl
Sequence 107, Appl
Sequence 108, Appl
Sequence 109, Appl

28 62.8 5.5 651 3 US-09-198-723A-92
29 62.8 5.5 651 3 US-09-198-723A-93
30 62.8 5.5 651 3 US-09-198-723A-94
31 62.8 5.5 651 3 US-09-198-723A-95
32 62.8 5.5 651 3 US-09-198-723A-96
33 62.8 5.5 651 3 US-09-198-723A-97
34 62.8 5.5 651 3 US-09-198-723A-98
35 62.8 5.5 651 3 US-09-198-723A-99
36 62.8 5.5 651 3 US-09-198-723A-100
37 62.8 5.5 651 3 US-09-198-723A-101
38 62.8 5.5 1998 3 US-09-198-723A-102
39 62.8 5.5 1998 3 US-09-198-723A-103
40 62.8 5.5 1998 3 US-09-198-723A-104
41 62.8 5.5 1998 3 US-09-198-723A-105
42 62.8 5.5 1998 3 US-09-198-723A-106
43 62.8 5.5 1998 3 US-09-198-723A-107
44 62.8 5.5 1998 3 US-09-198-723A-108
45 62.8 5.5 1998 3 US-09-198-723A-109

ALIGNMENTS

RESULT 1
US-09-094-435-3
; Sequence 3, Application US/09094435
; Patent No. 6251388
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
; TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN
; TITLE OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,435
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,085
; FILING DATE: June 9, 1997
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-094-435-3

Query Match 99.7% Score 1129.8; DB 3; Length 1133;

Thu Aug 28 09:00:59 2003

us-09-937-982-3.rni

802956 TACTGATGGATTGTCATTACCATGGTACAGATACCATGGAAGAGACGGCTATTTCCT 803015
426 CAACCTCAGGTGAAAGCCAAACCTGCTGCTGTAGGCGCCATGCTCCAGGCTC 485
803016 AGATTTAACCGTAAATGTGAAGAACCGGTGCTGCTGGGCAATGGCTCTCAAC 803075
486 TTCATGAGTGTGATGCCGCCATGATCTCTATAACCGCGTGAATGTAGGATCAACAA 545
803076 AGAAAGAGTGTGATGCCCATTAATCTTTACATGCTGTGCTGCGCAGCAGACAA 803135
546 AGCCTCTACTAAACAAAGAGTGTGATGTAGTAACGATGAGATTCACCGCCGACAGA 605
803136 AAAATCAAGTGTGCTGCTGCTTTAGTCGCAATGAATTAATGAAGTACTAGTGTCCGA 803195
606 AGCAGCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAAT 665
803196 TGTAAACAAACAGTAGAGCGGAGTGCAGTCCATTCACCAATATGCTTCT 803255
666 CGGCACAGTCTATTATGCAAGTGTGATTTTCACTCAATCCGTTGACCTCACACCT 725
803256 AGGCTATATTCAACAGCAAGTGGACTATGAACGTTCCCGAGAAAGCAACATACAT 803315
726 TGCAGTGTGATTTGATTAACAAATCGAAGACTCCCGAGAGTGCATATTTTACGC 785
803316 CAACACTCCGTTTAACTAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 803375
786 TCACCCGATGATGATGATGATTTTGTAGTCAATGCAAGTTCAGGAGGAGCAAGGAT 845
803376 TTAATCAATGACCTGTGCAACCAATTAAGCGATTAATCAATGCTGCTATCAAGGAT 803435
846 CATCATGAGGATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 905
803436 TGTATCTGAGGAGTGTGCAATGGAATGTTAAGTCTGCACACTTAGATGCTTAGAATA 803495
906 AGCAGCAAAATCAGCGGTGATGCTGCTCGAAGCTCTAGAGTGGGAGTGTTCACAC 965
803496 AGCCCAAAAGATAGCGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
966 CCAAGAGGCTGAAGTGTGATGATGAAGAACTTGGTGTGAGGCTAGAGAGTCTCAACC 1085
803556 TCGTACGCTGAAGTGTGATGATGAAGAACTTGGTGTGAGGCTAGAGAGTCTCAACC 1085
1026 TCAAAAGCCAGTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 803615
803616 ACAGAAAGCCGCTGCTTTCGATGATGATGATGATGATGATGATGATGATGATGAT 803675
1086 CCAAAAGATCTTCTCCACCTATTAA 1110
803676 TCAACATATTTGAGAGATTTCTAA 803700

RESULT 3
US-09-643-990A-1
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PE186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
Query Match 26.2%; Score 296.4; DB 4; Length 1830121;
Best Local Similarity 57.1%; Pred. No. 2.9e-86;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
QY 126 ACCCCAGTACTCTCTAGCCAGGAGGACCACTCGTGTTCGGGGGAATCTAGCCT 185
DB 802722 ACCAAATTAATCTTGGCAACGGGTGGTACCAATGCAGGAAGGCAAGTTCGT 802781
QY 186 CAAGAGTAGTACTCTCTGAGGAGTCAACCGTTGATAGCTTCTTGAGCGCTCCCTGC 245
DB 802782 AAATTCGCTATTAAGCTGGACATTAAGTATTGATATCTTTAATGAGCTGTACCA 802841
QY 246 CATCAACGACCTAGCCACCCTCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAT 305
DB 802842 AATGAAAATATTTGCCAATTAAGGTTGAGCAATTTGAAAATAGTTTCAAGACAT 802901
QY 306 GACGGTAAAGTGTGGCTTTAAACTAGCCAAAGCGTGTCAATAGCTCTCTCGCCCAAAA 365
DB 802902 GAATGACGAGTCTGGCTAAACTGGCAAAAGCCATCAATGCTC-----AATGTA 802955
QY 366 GACCGAAGCGGTGATCATCACCATGGAACTGACACCATGGAAGAGACCGCTTCTTCT 425
DB 802956 TACTGATGATTTGTCTATTCACCATGGTACAGATACCATGGAAGAGACCGCTTAT 803015
QY 426 CAACCTCAGGTGAAAGCCAAACCTGCTGCTGTAGGCGCCATGCTCCAGGCTC 485
DB 803016 AGATTTAACCGTAAATGTGAAGAACCGGTGCTGCTGGGCAATGGCTCTCAAC 803075
QY 486 TTCATGAGTGTGATGCCGCCATGATCTCTATAACCGCGTGAATGTAGGATCAACAA 545
DB 803076 AGAAAGAGTGTGATGCCCATTAATCTTTACATGCTGTGCTGCGCAGCAGACAA 803135
QY 546 AGCCTCTACTAAACAAAGAGTGTGATGTAGTAACGATGAGATTCACCGCCGACAGA 605
DB 803136 AAAATCAAGTGTGCTGCTGCTTTAGTCGCAATGAATTAATGAAGTACTAGTGTCCGA 803195
QY 606 AGCAGCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAAT 665
DB 803196 TGTAAACAAACAGTAGAGCGGAGTGCAGTCCATTCACCAATATGCTTCTCT 803255
QY 666 CGGCACAGTCTATTATGCAAGTGTGATTTTCACTCAATCCGTTGACCTCACACCT 725
DB 803256 AGGCTATATTCAACAGCAAGTGGACTATGAACGTTCCCGAGAAAGCAACATACAT 803315

QY 726 TCGAAGTGAGTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATCTTTACGC 785
Db 803316 CAACACTCCGTTTAACTAGAGAAAATTTAGATAGCTTACCACAAAGTGGGATTTATTCG 803375
QY 786 TCACCCCGATGATCTGATGTTTGTAGTCAATGAGCGCTTCAGCGAGGAGCCAAAGAAAT 845
Db 803376 TTATTCAAATGCACTGTGCAACCAATTAACGCAATTAATCAATGCTGGCTATCAAGGAT 803435
QY 846 CATCCATGAGGAGATGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTTTGAAA 905
Db 803436 TGTATCTGCAGAGTTGGCAATGGAATGTTAATGCTGCACATTTAGATCGCTTAGAAA 803495
QY 906 ACAGAGCAATCAGCGGTAGTCGCTCGAAGCTCTAGAGTGGCGAGTGGTTCCACCAC 965
Db 803496 AGCCGCAAAAGATAGCGTGGTGTAGTCCGTTCTCTCGTGTACCAACGGTTATACAC 803555
QY 966 CCAAGAGGCTGAAGTGGATGATAAGAACTTGGTTTTGTGGCTACAGAGAGTCTCAACC 1025
Db 803556 TCGTGACGCTGAAGTTGATAGTAAATATGGCTTTGTAGCATCAGGTACTTTAAATCC 803615
QY 1026 TCAAAAGCCAGAGTCTTTATGTTAGCCCTCACAAAACCTAGTGATAGAGAGCGAT 1085
Db 803616 ACAGAAAGCCCGTCTCTTCAATTTAGCTTAACTCAAACTAAGATCCNAAAGTAA 803675
QY 1086 CCAAGAGATCTTCTCCACCTATTA 1110
Db 803676 TCAACAATATTTGGAAGACTTCTAA 803700

RESULT 4

US-08-447-422-15
; Sequence 15, Application US/08447422
; Patent No. 5686579
; GENERAL INFORMATION:
; APPLICANT: SHAMI, Ezekiel Y.
; APPLICANT: ROTHSTEIN, Asar
; APPLICANT: RAMJESINGH, Mohabir
; TITLE OF INVENTION: Use of Antibody/Antigen Interactions To
; TITLE OF INVENTION: Protect or Modulate Biological Activity
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/447,422
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,410
; FILING DATE: 22-JUN-1993
; APPLICATION NUMBER: US 07/938,505
; FILING DATE: 31-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/205,748
; FILING DATE: 21-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17923/102 HYLI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-422-15

Query Match 25.9%; Score 294; DB 1; Length 1848;
Best Local Similarity 56.3%; Pred. No. 2.6e-87;
Matches 574; Conservative 0; Mismatches 440; Indels 6; Gaps 1;

QY 91 GGTGCGGATCCAGCGCTCTGTTTGTATGGTAAACCCCAAGTGACTTCTCTGAGCA 150
Db 835 GGTGTTGGTTCGGGTGGCGGATCTAGATCTTACCAATATACCAATTTTAGCAAC 894
QY 151 GGAGGACCATCGCTGGTTCGGGGGAATCTAGGCTCAAGAGTAGTCTCTCTGTGAGCA 210
Db 895 GCGGGACCATTTGCGGTGGTGTGACTCCGCAACCAATCTAATACAGTGGGTAA 954
QY 211 GTCACCGTTGATAAGCTTCTTGCAGCGTCCCTGCCATCAACGACCTAGCCACCATCAAG 270
Db 955 GTTGGCGTAGAAAATCTGTTAAATGCGGTGCGCAACTAAAGACATTCGCAACGTTAA 1014
QY 271 GGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGTAAAGTGTGGCTTAACTA 330
Db 1015 GCGAGCAGGTAGTGAATATCGCTCCCGAGACATGAACGATAATGTCTGGCTGACATG 1074
QY 331 GCCAAGCGTGAATGAGTCTCTCGCCCAAAAGAGACCGAGCGCTGATCATCACCAT 390
Db 1075 GCGAAAAAATTA-----CACCGACTGCGATAAGACCGCGCTTCGTCTATACCCAC 1128
QY 391 GGAATGACACCATGGAAGACACCGCTTCTTCTCAACCTCAGGTGAAAAGCCAAA 450
Db 1129 GGTACCGACAGTGAAGAAATCTGTTACTTCTCGACCTCGCGTGAATTCGACAAA 1188
QY 451 CCTGCTCTGCTTGTAGCGCCATCGCTCCAGCTTCTTCATAGTGTGTGAGGCCCCATG 510
Db 1189 CCGGTGTTGATGTTGCGGCAATGCGTCCGTCACAGTCTATAGCGGACAGCGTTCATTC 1248
QY 511 AATCTCTAATACCGCTGAATGTAGCATCAACAAAGCTCTACTAAACAAAGAGTGTGTG 570
Db 1249 AACCTGTATACCGCGTGTAGTGAACGCTGATAAGCTTCGCAACCGCTGCGGTGTGTG 1308
QY 571 ATTGTGATGAACGATGAGATTTCACGCGCGCAGAGAGCGAGCTCAACACACCGCA 630
Db 1309 GTAGTGAATGACACCGTCTTGTAGTGGCGGTGACGTCAACAAAACACACACCGCAC 1368
QY 631 GTCATGCAATTTGCTTGGCCCAACACAGGTAAATCGGCAAGTCTTATTTATGCAAGTC 690
Db 1369 GTAGCGACCTTCAAGTCTGTAACTAGCGTCTCTGGTTTACATTCACACGGTAAATTT 1428
QY 691 GAGTATTTCACTCAATCCGTTTCGACCTCAGCTTGCACCTTGCAGAGTGTGTATATGCAAA 750
Db 1429 GACTACAGGTTACCCCGGACGTAAGCATACCGGACACCGCATTCGATCTCTTAAG 1488
QY 751 ATCGAAGAACTCCCGAGAGTCAATTTCTTTCAGCTCACCCCGATGATATCTGATTTTA 810
Db 1489 CTGAATGAATGCGGAAAGTGGCATTTTAACTACGCTAACGATCCGATCTTCG 1548
QY 811 GTCATGCAAGCTTTCAGGAGAGGCAAGGAATCATCATGACAGGATGGGCAATGGG 870
Db 1549 GCTAAGCACTGGTAGATCGGGCTATGATGGCATCGTTAGCGCTGGTGGGTACGGC 1608
QY 871 AACCCCTTTCCTTTGACTCAAAATGCTTTTGAAGAGCAGCAAAATCAGCGGTAGTCGTC 930
Db 1609 AACCTGTATAATCTGTGTTGCACAGCTGGGACCGCCGGAACCCGTAAGTCAAGTC 1668
QY 931 GCTCGAAGCTCTAGAGTGGGAGTGGTTCACACCCAGAGAGGCTGAAGTGAATGAAG 990
Db 1669 GTGCGTCTTCTCCCGGTACCGACGGGCGTACCACTCAGGATGCCGAGTGGATGATCG 1728
QY 991 AAACTTGGTTTGTGCTACAGAGTCTCAACCTCAAAAGCCAGAGTGTCTTCTATG 1050

Thu Aug 28 09:00:59 2003

Db 1729 AAATACGGCTTCGTGCTCTGCGACGCTGAAACCGGCAAAAGCGCGCTTCTGCTGCAA 1788
Qy 1051 TTAGCCCTCACCAAACTAGTATGATAGAGCGGATCCAAAAGATCTTCTCCACCTATTAA 1110
Db 1789 CTGGCTCTGACGCAAAACCAAGATCCGACGAGATCCAGCAGATTTCAATCAGTAA 1848

RESULT 5
PCT-US92-10421-1
; Sequence 1, Application PC/TUS9210421
; GENERAL INFORMATION:
; APPLICANT: Roberts, Joseph
; APPLICANT: MacAllister, Thomas W.
; APPLICANT: Sethuraman, Natarajan
; APPLICANT: Freeman, Abbie G.
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
; TITLE OF INVENTION: ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie and Beckett
; STREET: 1001 G Street N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20001

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10421
; FILING DATE: 19921204
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: DE P4140003.8
; FILING DATE: 04-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 00100.41200
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1014 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: PSEUDOMONAS 7A
; PCT-US92-10421-1

Query Match 19.0%; Score 215.8; DB 5; Length 1014;
Best Local Similarity 52.8%; Pred. No. 1.9e-61;
Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;
Qy 128 CCCAAGTGACTATCTTAGCCACAGGAGGACCAATCCCTCGTTCGGGGGAATCTAGGTCA 187
Db 29 CCAAGTGGTGAATCTTGGCCACCGCGGACCATCCCGCGCTGCGCCAGCGGGCA 88
Qy 188 AGAG---TAGCTACTCTGTGGAGAGTACCGTTTATTAAGCTTCTTGAGCGGTCCTG 244
Db 89 ACAGCGCCACTACAGGCTGCCAAGGTTGGCGTGCACAGCTGATGGCGGGTCCCG 148
Qy 245 CCATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGA 304
Db 149 AGCTGGCGGACCTGGCCCAATGTGGCGGCGAGGATGATGACATCGCTCCGAAAGCA 208
Qy 305 TGACGGGTAAGGTGTGGCTTAAACTAGCCAAAGCGTGTCAATGAGCTCTCGCCCAAAAG 364

Db 209 TCACCAACGACGACCTGCTCAAGCTGGCAAGAGCGTGGCCGAGCTGGCCGACAGCAATG 268
Qy 365 AGACCGAAGCGGTGATCATCAACCATGGAATCAGACCAATGGAAGAGACGCTTTCTTCC 424
Db 269 ACGTCGATGCGATCGTTCATCAACCATGGAACCGGACCTTGGAAAGAACCGCTACTTTT 328
Qy 425 TCACCTCAGCTGAGAAAGCAAAACCTGCTGCTCTTGTAGGCGGCATCGCTCCAGCT 484
Db 329 TGAACCTCGTGGAAAGACCGCAAGCCGATCGTGGTTCGTTCCATGCGCCCGCA 388
Qy 485 CTTCCATGAGTGTGATGGCCCATGATCTTATACCGCGGTGAATAGCGATCAACA 544
Db 389 CCGCATGTTCGGCGACGGCATGTCAACCTGTACAACGCGCTGGCCGTGGCCGACCA 448
Qy 545 AAGCTCTACTAAACAAAGGAGTGTGATTTGATGAACGATGAGTTCACGCCCCAGAG 604
Db 449 AGGACTCGCGCGCAAGGGCGTGTGTGACCATGAACGACGAGATCCAGTCCGGGGTG 508
Qy 605 AAGGACCAAGCTCAACACCAACCGCAGTCAATGATTTGTTCCGCCCAACACAGTAA 664
Db 509 ACGTGAGCAAGTCGATCAACATCAAGACCGAAGCTTC---AAGAGCGCTTGGGGCCGC 565
Qy 665 TCGGCACAGTCTATTATGCGCAAGTCGAGTATTTCACTCAATCCGTTCCGACTCACACCC 724
Db 566 TGGGCAATGGTGGTGAAGGCAAGTGTACTGTTCCGCTCGCGCCGCGCAAGCCACACGG 625
Qy 725 TTGCAAGTGAATTCATATAGCAAAATCGAAGAACTCCCGAGAGTCGATATTTTACG 784
Db 626 TCAATCCGAGTTCGACATCAAGCAGATCAGCAGCTGCCCGAGGTGACATCGCTACA 685
Qy 785 CTCACCCCGATGATCTGATGTTTGTAGTCAATGAGCCCTTCAGGACGAGGCAAGGAA 844
Db 686 GCTATGGCAACGCTCACACACCGGCTTACAGGCCCTGGCACAGAACGCGCAAGCGC 745
Qy 845 TCATCCATGCGAGCATGGCAATGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGA 904
Db 746 TGATCCATCCCGCACCGCAATGCTCGGTCTCGTGGGGTGGTGGCCGCTGCAGG 805
Qy 905 AAGCAGCCAAATCAGCGGTAGTGTGCTGCGTCAAGCTCTAGAGTGGGAGTCTTCCACCA 964
Db 806 AGCTGCGCAAGAACGCGCTGCAGATCATTCGTTCTGTCACGTCAACAGGCGGTTCTG 865
Qy 965 CCCAAGAGGCTGAAGTGAATGATAGAAACTTGGTTTGTGCTACAGAGAGTCTCAACC 1024
Db 866 TCGTAAACCGCGAGCGCCGACGACGACGACGACGACTGGGTGCTGGCCACGACCTGA 925
Qy 1025 CTCAAAAAGCCAGAGTGTCTTTATGTTAGCCCTTACCAAACTAGTGTAGAGAGCGCA 1084
Db 926 CGCAGAAAGCCCGCATCTCTGGCGATGGTGGCAATGACCAAGACCCAGGACGAGGAGC 985
Qy 1085 TCCAAAAGATCTTCTCCACCTATAAT 1111
Db 986 TGCAGCGCATTTTCTGGGAATACTGAT 1012

RESULT 6
US-08-050-482A-1
; Sequence 1, Application US/08050482A
; Patent No. 6312939
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, Joseph
; APPLICANT: MACALLISTER, Thomas W.
; APPLICANT: SETHURAMAN, Natarajan
; APPLICANT: FREEMAN, Abbie G.
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
; TITLE OF INVENTION: ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.

COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/050,482A
 FILING DATE: 25-Apr-1995
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US92/10421
 FILING DATE: 04-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 023032/0106
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1017 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1011
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-08-050-482A-1

Query Match
 Best Local Similarity 18.6%; Score 210.8; DB 4; Length 1017;
 Matches 524; Conservative 0; Mismatches 457; Indels 9; Gaps 3;
 128 CCCAAGTACATCTCTAGCCACAGGAGGACCATCGCTGGTTCGGGGAATCTAGCGTCA 187
 29 CCAAGTGTGATCTCTGCCACCGCGGCACATCGCGCGCTGGCGCAGCGCGGCA 88
 188 AGAG---TAGTACTCTGTGGAGCAGTCACTGTTGATTAAGCTTCTTGACCGCTCCCTG 244
 89 ACAGCGCCACCTACAGGCTGCCAAGGTTGCGCTCGACAGCTGATTCGCGCGCTGCGG 149
 245 CCATCAACGACCTAGCCACCATCAAGGTTGAGTGAACAGATCTCAAGCATTTGGTCCCAAGAGA 304
 149 AGCTGGCGGACCTGGCCAAATGTGCGCGCGAGCAGAGTGATGCAGATCGCTCCGAAAGCA 208
 305 TGACGGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATGAGTCTCTCGCCCAAAAG 364
 209 TCACCAACGACGACCTCTCAGCTGGGCAAGCGGTGGCGGAGCTGGCGCAGCAATG 268
 365 AGACCGAGCGGTGATCATCCCATGGAATCAACCATGGAAGACCGCTTTCTTTC 424
 269 ACCTGATGGCATCGTCTACCCCATGGCACCCGACACCTTGGGAAGAACCGCTACTTTT 328
 425 TCAACTCTACGGTGAAGGACCAAAACCTGTCTGCTTGTAGCGGCATCGTCCAGGCT 484
 329 TGAACCTCTGGAAAGACGACAGCCGATCTCGTGTGCTGCTCCATGCGCGCGGCA 388
 485 CTTCCATGATGTGATGCGCCCATGAATCTTATAACCGCGTGAATGATGAGCATCAACA 544
 389 CGGCCATGTCCGCGGACGGCATGCTCAACTGTATCAACGCGGTGGCGTGGCGCAGCAACA 448
 545 AAGCCCTCTTAAACAGAGGTGATGATTTGATGATGATGATGATGATGATGATGATGATGAT 604
 449 AGGACTCGCGGCAAGGCGGTGCTGGTGAACATGAACGACGAGATCCAGTCCGCGGTG 508
 605 AAGGACCAAGTCAACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAA 664
 509 ACGTGAGCAAGTCAATCAATCAAGACCGAAGCCTTC---AAGAGCGCTGGGCGCGC 565

665 TCGGCAAGTCTATTATGCGAAAGTCGAGTATTCTCACTCAATCGTTTCGACCTCACACCC 724
 566 TGGGCAATGTTGGTGGAGGCAAGTCTGACTGTTGCTCGGCGGCAAGCGCCACACGG 625
 725 TTGCAAGTCAAGTTGATATAGCAAAATCGAAGAACTCCCCAGAGTGAATATTTTTCAG 784
 626 TCACTCCGAGTTCGACATCAAGCAGATCAGCAGCCTGCCAGGTTGGACATCGCCTACA 685
 785 CTCACCCCGATGATGATGATGTTTATGTCATCAATGACAGCCCTTCAGGAGGAGCCAAAGGAA 844
 686 GCTATGGCAACGTCCAGCAGCGCTCAAGCGCTTGGCAAGAGCGCGCGCAAGCGCG 745
 845 TCATCATCAGGAGTGGCAATGGGAACCTTTTCCCTTTTGAATCAAAATGCTCTTGA 904
 746 TGATCATGCGCGCACCGGCAATGCTGCTGCTGCGGGTGGTGGCAGCCCTCAGG 805
 905 AAGCAGCCAAATCAGGCGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
 806 AGCTGCGCAAGAACCGCGGTGACATCATTCGTTCTGCTCCACGTCACCAAGCGCGGTTTCG 865
 962 CCACCAAGAGGCTGAAGTGGATGATAAGAACTTTGTTTGTGCTACAGAGTCTCA 1021
 866 TGCTGCTGAACCGCGAGCGCCGACGACAGCAAGCAAGTGGTCTGGCGCCACGACCTGA 925
 1022 ACCCTCAAAAAGCCAGAGTCTTTTATGTTAGCCCTCACCAGAACTAGTATGATAGAGG 1081
 926 ACCCGCAGAGCGCGCATCTCTGGCGATGTTGGCAATGACCAAGACCCAGGACAGCAAGG 985
 1082 CGATCCAAAAGATCTTCTCCACCTATTAAAT 1111
 986 AGCTGCGAGCGCATTTTCTGGGAATCTGAT 1015

RESULT 7
 US-09-252-991A-5793
 ; Sequence 5793, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5793
 ; LENGTH: 1182
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5793

Query Match
 Best Local Similarity 17.7%; Score 200.4; DB 4; Length 1182;
 Matches 513; Conservative 0; Mismatches 456; Indels 9; Gaps 3;
 128 CCCAAGTACATCTCTAGCCACAGGAGGACCATCGCTGGTTCGGGGAATCTAGCGTCA 187
 197 CCAAGTGTGATCTCTGGCGCACCGCGGCACCATCGCGCGCGCGCGCGCGCGCGCGCA 256
 188 AGAG---TAGTACTCTGTGGAGCAGTCAACCGCTGATGATGATGATGATGATGATGATGAT 244
 257 ACAGCGCCACCTACACCGCGCCAGGTTGCGCGTGCACAGTGTGTCGCGCGCGCGCGCG 316
 245 CCATCAACGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGA 304
 317 AGCTCAAGGACATCGCCCAACGTCGCGCGCAACAGGTGTTCCAGATCGCTCGGAAAGCT 376
 305 TGACGGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATGAGTCTCTCGCCCAAAAG 364

QY 1082 CGATCCAAAGATCTTCT 1099
 DB 254 AGCTGCAACGGATCTTCT 237

RESULT 9

US-09-328-352-2336
 ; Sequence 2336, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2336
 ; LENGTH: 1125
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-2336

Query Match 14.1%; Score 159.2; DB 4; Length 1125;
 Best Local Similarity 50.3%; Pred. No. 1.4e-42;
 Matches 504; Conservative 0; Mismatches 483; Indels 15; Gaps 4;
 QY 121 GCTAAACCCCAAGTACTATCTAGCCACAGGAGCCACCATCGCTGGTTCGGGGAAATCT 180
 DB 127 GCAAAACATGTGTGTGCTACTGCGGTACCATGCTGGTGTGGCGCAAGC 186
 QY 181 AGCTCAAGAGT---AGTACTCTGCTGGAGCAGTCAACGTTGATAAGCTTCTTGACGCC 237
 DB 187 TCAGCAATAGTGAACATTATACGGCGCAAAAGTTCCAGTTGATGCTTTAATCAATGCA 246
 QY 238 GTCCCTGCTCAACGACTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGCTCC 297
 DB 247 GTTCTCAATTCAGATTGGCGAATGTATCTGGTATCCAAAGCATTAAGTAGCTTCT 306
 QY 298 CAAGAGATGACGGGTAGGTGTGGCTTAACTAGCAAGCGGTGTCAATGAGCTCTCGCC 357
 DB 307 GAAAGTATTACTGACAGGAATATTACAAATGCTCGTCAAGTAATGAATCTTTAA 366
 QY 358 CAAAGAGACCGAAGCGGTGATCATACCCATGGAAGTACACCATGGAAGAGACCGCT 417
 DB 367 AAGCAACTGTAAATGGCGTTGTGATTACACAGGTACAGATATCTTTAGAGAAACAGCA 426
 QY 418 TTCTTCTCAACCTCACGGTGAAAGCCAAACCTGTCTGCTTGTAGGCGCCATGCT 477
 DB 427 TTTTCTTAATCTTGTGTTCACTGATTAACCAATCGTACTTGTGGTCAATGCGC 486
 QY 478 CCAGGCTCTTCCATGAGTGTGATGGCCCAATGAATCTTATAAGCGCGTGAATGTAGCG 537
 DB 487 CCATCAACTGCTCTTTCAGCAGATGTTCCACTTAACCTTTATAGTGTGTTGCAATGGCC 546
 QY 538 ATCAACAAAGCCCTACTAACAAGAGAGTGTGATTTGATGAACGATGAGATTAACGCC 597
 DB 547 GCTTCTGATGATGCAAAAATAAAGGCGTTATGTTCTCATGAACGACTCTATTTTGTCT 606
 QY 598 GCCAGAGAACGACCAAGCTCAACACCAACCGCAGTCAATGATGATTTGCTTCGCCCAACACA 657
 DB 607 GCTCGTATGTAATAAGGCAATTAACTATACAAATGCTTTTGAAGCC---AATGG 663
 QY 658 GGTAAATCGGCAAGCTTATATGGAAGTGAAGTATTTCACTCAATCGTTTCGACCT 717
 DB 664 GGTGCTTTGGGTACACTTGTGAAGGCAACCAATTTGGTTTGAACAATCTGTTAAAGCT 723
 QY 718 CACACCTTTCGAAGTGTGATAT-----TAGCAAAATCGAAGTATCCCGCAGATC 771
 DB 724 CATACCAATGCTTCAGAAATTAATTTGAATAATTAAGGATGATGCTCTTCGACAGTA 783
 QY 772 GATATTCTTTACGCTCACCCCGATGATGATGATGTTTGTAGTCTTTCAGTCAATGAGCGCTTCAGGCA 831

DB 784 CAAATCGTTTATGTTCTGACTCTATGCTTCTGATGCTTATGAAGCATATGCAAAAGCT 843
 QY 832 GGAGCCAAAGGAATCATCCATGAGGATGGCAATGGAA---CCCTTTTCCCTTTGACT 888
 DB 844 GCGCTAAAGCGATTATTCATGCAAGTACAGGTAATGGTTCTGTAGCGAAATATATCGTT 903
 QY 889 CAAATGCTCTTGAAGAAGCAGCAAAATCAGGCGTAGTCGCTCGAGCTCTTAGAGTG 948
 DB 904 CCAACATTACAAACCTTCATGACAAAACGGTATTCAAAATATATCCGCTCATCGCGCTT 963
 QY 949 GGCAGTGGTTCACCAACCCCAAGAGCTCAAGTGCATGATAGAAACTTTGTTTCTGGCT 1008
 DB 964 CCACAAGTTTGTATTACGTGATGAGAACCACTGACTTAATATATGTTGGTAGCA 1023
 QY 1009 ACAGAGATCTCAACCTCAAAAGCCAGAGTGCTTTTATGTTAGCCCTCACCACAACT 1068
 DB 1024 GCTCATGATTGAATCTCAAAAAGCGCTCTCTTGGCGGTTGGCACTTACCACAAACC 1083
 QY 1069 AGTGATAGAGAGCGGATCCAAAGATCTTCTCCACCTATTAA 1110
 DB 1084 AACGATGCGAAGAAATTCAACGATGTTCTGGCAGTACTAA 1125

RESULT 10

US-09-252-991A-5773
 ; Sequence 5773, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5773
 ; LENGTH: 2748
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5773

Query Match 13.0%; Score 147.4; DB 4; Length 2748;
 Best Local Similarity 51.8%; Pred. No. 2e-38;
 Matches 384; Conservative 0; Mismatches 351; Indels 6; Gaps 2;
 QY 362 AAGAGACCGAAGCGCTGATCATCCCATGGAACCTGATGAGAGAGAGACCGCTTCT 421
 DB 3 ACGAGCTGACGGCATCGTATCACCCAGGTACCGACACCTTGAAGAGACCGCTTCT 62
 QY 422 TCCTCAACCTCAGGTGAAAAGCCAAAACCTGTCTGCTTGTAGCGCCATCGTCCAG 481
 DB 63 TCCTGACCTGTGTCAGACACCGAGAGCTTATCGTGTGTCGCTCGATGCGCCGG 122
 QY 482 GCTCTTCCATGAGTCTGATGCGCCCATGAATCTCTATAACGCGGTGAATGTAGCATCA 541
 DB 123 GCACCGCATGTCGCCCGCAGCGCATGCTCAACCTGTATCAACCGCTGGCGGTGGCGG 182
 QY 542 ACAAGAGCTCTACTAACAAAGAGTGGTGTGATGAACCATGATGAGATTACGCGCCCA 601
 DB 183 ACAAGTGGCAGCGGCAAGGCGTGTGATCACCATGACGAGAGATCTCTCGGCC 242
 QY 602 GAGAAGGACCAAGCTCAACACCGCAGTCAATGCTTGTGCTTGCCTGCGCAACACAGGTA 661
 DB 243 GCGACGCGAGCAAGATGCTCAACATCAAGACCGAAGCGTT---CAAGAGCCCGTGGGCGC 299
 QY 662 AAATCGGCACAGTCTTATTATGCAAGTCCGATTTTCACTCAATCGCTTCGACCTCACA 721
 DB 300 CGCTGGCATGCTGTGTCGAGGGCAAGAGCTACTGTTTCCGCGCACCGGTGAAGCGGCACA 359

	Query Match	8.9%;	Score 101.4;	DB 4;	Length 552;
	Best Local Similarity	84.4%;	Pred. NO. 1.5e-23;		
	Matches 114;	Conservative 0;	Mismatches 21;	Indels 0;	Gaps 0
Qy	1	ATGGGCAGCAGCCATCATCATCATAGCAGCGCCTGTGTCGCGCGGCAGCCAT	60		
Db	1	ATGGGCAGCAGCCATCATCATCATACAGCAGCGCCTGTGTCGCGCGGCAGCCAT	60		
Qy	61	ATGGCTAGCATGACTGTGGACAGCAATGGGTTCGCGGATCCAGGCGCTCTCTTTTGATG	120		
Db	61	ATGGCTAGCATGACTGTGGACAGCAATGGGTTCGCGGATGGGAATTCGCTGGCTTTGTT	120		
Qy	121	GCTAAACCCCAAGTG	135		
Db	121	GGTAACAAAGCAGTG	135		

RESULT 13
US-09-107-532A-1293
; Sequence 1293, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1293:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...1005
; SEQUENCE DESCRIPTION: SEQ ID NO: 1293:
US-09-107-532A-1293

```

```

Query Match      6.8%; Score 76.6; DB 4; Length 1005;
Best Local Similarity 53.5%; Pred. No. 3.9e-15;
Matches 160; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 376 GTGATCATACCCATGGAACTGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCAGC 435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 GTGGTGATCACACACCGGTACAGATACATTAGAGAACTGCTTTTCTTGTATACACG 336
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GTGAAGCCAAACCTGCTGCTGTAGCGCCATGCGTCCAGGCTCTTCCATGAGT 495
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 ATTGGCAACAAATTACCAATTTGTTGACTGGGCAATGCGTTCAAGTAATGAATTGGA 396
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GCTGATGGCCCCCATGAATCTCTATAAGCGCGTGATGTAGCGATCAACAAGCCTCTACT 555
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 AGTACGGCCTTTAATTTTGAAGTGCAGTTCTGTGTCGAAGTTGTGAAGAACGCGTA 456
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 AACAAAGAGTGTGATGTGATGAACGATGAGATTCACGCGCCGAGAGAGCCACCAAG 615
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 GATAAGGGGGTTTGGTGGTCATGAACGATGAGATCCATTCTGACGCTACGTGACAAAG 516
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 CTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGATAAAATCGGCACAGT 674
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 ACCATACCAAAATGTTGCTACTTTCCGAAACCTTACATTGGGTGCCATCGGCTTAGT 575
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-08-961-527-265
; Sequence 265, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 265:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-265

Query Match      6.5%; Score 74; DB 4; Length 2231;
Best Local Similarity 51.9%; Pred. No. 4.7e-14;
Matches 190; Conservative 0; Mismatches 175; Indels 1; Gaps 1;

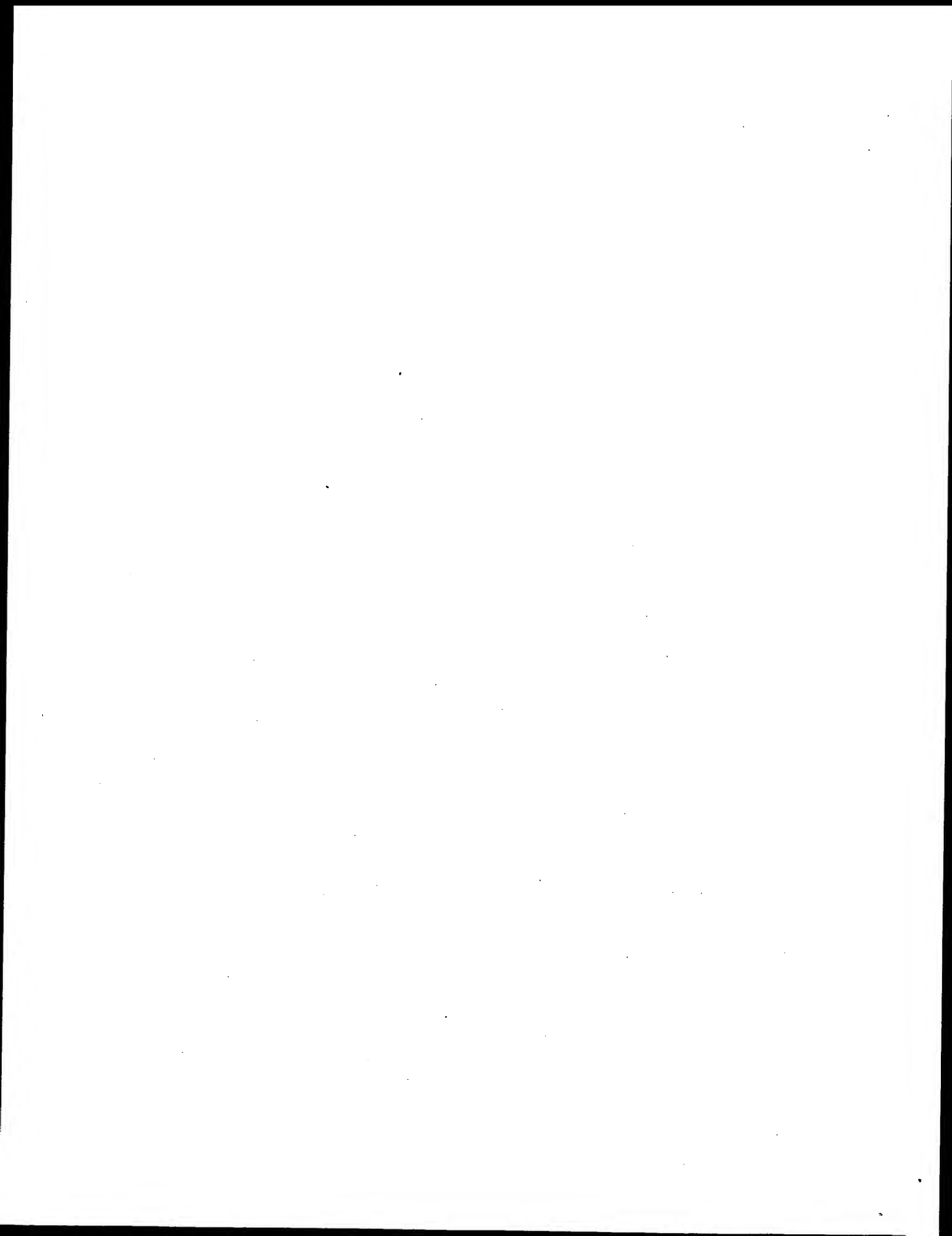
QY 369 CGAAGCCGTGATCATCACCCATGGAACCTGACACCATGGAAGAGACCGCTTTCTTCTCAA 428
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 429 CCTCAGCGTGAAGCCAAACCTGCTGCTGTAGGCGCCATGCGTCCAGGCTCTTC 488
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1137 TACCATGGAAGTTCCCATATGCTTCTTAACAGGAGCCATGGTACTC-CAATGA 1195
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 CATGAGTGTGATGGCCCCCATGAATCTCTATAAGCGCGTGAATGTAGGATCAACAAGC 548
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1196 GCTCGGTAGTGTGCTTTTATTAATTAACCTTAAGTGTCTTAGAGTGGCCAGCGATGACAG 1255
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 CTCTACTAAACAAAGAGTGTGATGTGATGAACGATGAGATTCACGCGCCGACAGAGAGC 608
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1256 GGCTGCTGCAAAAGAGTGTGCTGTTATGAACGATGAATCCACGCTGCCAAGTATGT 1315
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 GACCAGCTCAACACACACCGCAGTCAATGCTTTCGCCCAACACAGAGTAAATCGG 668
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1316 CACCAAAACACATACGACTAATGTGACGACTTCCAGACTCCAAACATGCCCCCTTGG 1375
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 CACAGTCTATTATGGCAAGTCGAGTATTTCATCAATCGTTCCGACTCACACCCCTTGC 728
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1376 TCTCATCAUGAAACAGGAATCCTTACTTCAAAACAGCTGAACCTCGTGTGCTTTGA 1435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 AAGTGA 734
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1436 CCTTGA 1441
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-613-303-11
; Sequence 11, Application US/09613303
; Patent No. 6495347
; GENERAL INFORMATION:
; APPLICANT: Siegel, Marvin
; APPLICANT: Chu, N. Randall
; APPLICANT: Mizzen, Lee A.
; TITLE OF INVENTION: INDUCTION OF A TH1-LIKE RESPONSE IN VITRO
; FILE REFERENCE: 12071/002001
; CURRENT APPLICATION NUMBER: US/09/613,303
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: US 60/143,757
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion sequence
; NAME/KEY: CDS

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 02:34:35 ; Search time 843.661 Seconds
(without alignments)
3081.078 Million cell updates/sec

Title: US-09-937-982-3
Perfect score: 1133
Sequence: 1 atgggcagcagccatcatca.....asgaagggaatcttctcac 1133

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubnpa/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubnpa/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubnpa/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubnpa/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubnpa/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubnpa/US08_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubnpa/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubnpa/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubnpa/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubnpa/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubnpa/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubnpa/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubnpa/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubnpa/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubnpa/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubnpa/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubnpa/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	10	US-09-773-260-3
2	296.4	26.2	1830121	14	US-10-329-360-1
3	259.2	22.9	1196	12	US-09-882-227-83
4	215.8	19.0	1014	9	US-09-842-628-1
5	103.6	9.1	543	10	US-09-876-348A-34
6	103.6	9.1	543	10	US-09-876-796A-34
7	102	9.0	543	10	US-09-876-348A-18
8	102	9.0	543	10	US-09-876-348A-22
9	102	9.0	543	10	US-09-876-348A-30
10	102	9.0	543	10	US-09-876-796A-18
11	102	9.0	543	10	US-09-876-796A-22
12	102	9.0	543	10	US-09-876-796A-30
13	102	9.0	543	10	US-09-876-796A-34
14	102	9.0	543	10	US-09-876-796A-38
15	101.6	9.0	1436	12	US-10-228-063-43
16	100.4	8.9	681	10	US-09-876-348A-16

17	100.4	8.9	681	10	US-09-876-348A-28	Sequence 28, Appl
18	100.4	8.9	681	10	US-09-876-348A-36	Sequence 36, Appl
19	100.4	8.9	681	10	US-09-876-796A-16	Sequence 16, Appl
20	100.4	8.9	681	10	US-09-876-796A-28	Sequence 28, Appl
21	100.4	8.9	681	10	US-09-876-796A-36	Sequence 36, Appl
22	100.4	8.9	682	10	US-09-876-348A-20	Sequence 20, Appl
23	100.4	8.9	682	10	US-09-876-348A-32	Sequence 32, Appl
24	100.4	8.9	682	10	US-09-876-796A-20	Sequence 20, Appl
25	100.4	8.9	682	10	US-09-876-796A-32	Sequence 32, Appl
26	100.4	8.9	682	10	US-09-876-796A-24	Sequence 24, Appl
27	100.4	8.9	776	10	US-09-876-796A-24	Sequence 24, Appl
28	98.4	8.7	1435	12	US-10-228-063-41	Sequence 26, Appl
29	96.4	8.5	543	10	US-09-876-348A-26	Sequence 26, Appl
30	96.4	8.5	543	10	US-09-876-796A-26	Sequence 26, Appl
31	94.2	8.3	900	10	US-09-850-677A-1	Sequence 1, Appl
32	93.4	8.2	5312	12	US-10-263-103-35	Sequence 35, Appl
33	72.6	6.4	366	14	US-10-267-311-11	Sequence 11, Appl
34	70.4	6.2	1230	14	US-10-267-311-54	Sequence 54, Appl
35	68.4	6.0	6682	10	US-09-070-927A-56	Sequence 56, Appl
36	66.2	5.8	2847	14	US-10-267-311-20	Sequence 20, Appl
37	65.6	5.8	6061	12	US-10-314-861-10	Sequence 10, Appl
38	65.6	5.8	6100	12	US-10-314-861-36	Sequence 36, Appl
39	65.6	5.8	6115	12	US-10-314-861-34	Sequence 34, Appl
40	65.6	5.8	6121	12	US-10-314-861-38	Sequence 38, Appl
41	65.6	5.8	6133	12	US-10-314-861-32	Sequence 32, Appl
42	65.6	5.8	6151	12	US-10-314-861-30	Sequence 30, Appl
43	65.6	5.8	6169	12	US-10-314-861-28	Sequence 28, Appl
44	65.6	5.8	6248	12	US-10-314-861-15	Sequence 15, Appl
45	65.2	5.8	1989	14	US-10-269-557-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-773-260-3
; Sequence 3, Application US/09773260
; Patent No. US20020102251A1
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
; ASPARAGINASE IN THE TREATMENT OF HUMAN
; HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/773,260
; FILING DATE: 31-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/094,435
; FILING DATE: 1998-06-09
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600


```

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c

```


Y 666 CGGCACAGTCTATTATGGCAAAGTCGAGTATTTCACCTCAATCCCTTCACCTCCACACCGCG

Db	697	AGGAGCGTGTATTTATGCAAAACGCGCTATTATCATGAGCCTTTGAGAAAACACACAC	756
Qy	726	TGCAAGTGAAGTTGATATTAGCAAAATCGAA---GAACTCCCGAGAGTCGATATTTCTTTA	782
Db	757	AGAGCGCAATTTTCCCTTTTACAACTCAAAACCCCTGCTTAAAGTGGATATTATTTA	816
Qy	783	CGCTCACCCGATGATGATGATTTAGTCAATGAGCCCTTCAGGAGGAGCAAGG	842
Db	817	CACGCATGCTGGCATGACCCCTGATTTATTCAGAGCGAGCCTAAATCGCATGCAAAAGG	876
Qy	843	AATCATCCATGACGAGTGGCAATGGAACCTTTCCCTTTGACTCAAAATGCTCTTTGA	902
Db	877	CGTTGTGATAGCCGGGTGGTAAATGGAAATGAGCGCTGGTTTTAAAGCGATGCA	936
Qy	903	AAAAGCAGCCAAATCAGCGGTAGTGTGCTGCGTCAAGTCTTAGAGTGGGAGTGTTCAC	962
Db	937	AGAAGCGAGCCAAATGGGGGTGGTATTTGTTCTTAGCAGGGGTAAATAGCGGTGAGAT	996
Qy	963	CACCAAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGTGGCTACAGAGAGTCTCAA	1022
Db	997	TA-----CTTCAGCGGAGATTGATGACAAGGCCCTTATCAAGCGACATTTAAA	1047
Qy	1023	CCCTCAAAAGCCAGAGTGTCTTTATGTTAGTCCCTCACCATACTAGTGTAGAGAGGC	1082
Db	1048	CCCCAAAAGCTAGGTGCTTTTACAACTCGCTTTAACTAAACAAATAATAAAGAAA	1107
Qy	1083	GATCAAAAGATCTTCTCCACTATTAA	1110
Db	1108	AATCCAAGAAATGTTTGAAGAGTATTGA	1135

RESULT 4
US-09-842-628-1
; Sequence 1, Application US/09842628
; Patent No. US20020064862A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: MACALLISTER, THOMAS W.
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: FREEMAN, ABBIE G.
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN
; TITLE OF INVENTION: ANTIVIRAL AND ANTICANCER THERAPY
; FILE REFERENCE: 023032/0108
; CURRENT APPLICATION NUMBER: US/09/842,628
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 08/050,482
; PRIOR FILING DATE: 1995-04-25
; PRIOR APPLICATION NUMBER: PCT/US92/10421
; PRIOR FILING DATE: 1992-12-04
; PRIOR APPLICATION NUMBER: DE P 4140003.8
; PRIOR FILING DATE: 1991-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1008)
US-09-842-628-1

Query Match
Best Local Similarity 52.8%; Score 215.8; DB 9; Length 1014;
Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;
Qy 128 CCCAAGTGACTATCTTAGCCACAGGAGGACCACTCGCTGGTTCGGGGGAATCTTAGCGTCA 187
Db 29 CCNACGTGGTATCTTGGCCACCGCGGACCACTCGCGCGCTGGCGGAGCGCGCA 88
Qy 188 AGAG---TAGCTACTCTGTGGAGAGTACCCCTTGATAGCTTTCTTGAGCGGTCCTG 244

[illegible]

RESULT 9						
US-09-876-348A-30						
; Sequence 30, Application US/09876348A						
; Patent No. US20020172951A1						
; GENERAL INFORMATION:						
; APPLICANT: Horwath, K. L. and Myers, K. L.						
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio						
; FILE REFERENCE: RB-125-R1						
; CURRENT APPLICATION NUMBER: US/09/876,348A						
; PRIOR FILING DATE: 2001-08-09						
; PRIOR FILING DATE: 2000-06-08						
; SOFTWARE: Microsoft Word						
; SEQ ID NO 30						
; LENGTH: 543						
; TYPE: DNA						
; ORGANISM: Tenebrio molitor						
; OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4						
US-09-876-348A-30						
Query Match 9.0%; Score 102; DB 10; Length 543;						
Best Local Similarity 88.1%; Pred. No. 3.8e-24;						
Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;						
QY	1	ATGGGCAGCAGCCCATCATCATCATCATCATGACGCGGCCCTGTGTCGCGCGGCGGCGGCCAT	60			
Db	64	ATGGGCAGCAGCCCATCATCATCATCATCATGACGCGGCCCTGTGTCGCGCGGCGGCGGCCAT	123			
QY	61	ATGGCTAGCATGACTGGTGGACAGCAATGGTGCGCGATCCAGCGCCTCTGTTTGTATG	120			
Db	124	ATGGTAGCATGACTGGTGGACAGCAATGGTGCGCGATCCCTCACCGAACAGATA	183			
QY	121	GCTAAA 126				
Db	184	CAGAAA 189				
RESULT 10						
US-09-876-348A-38						
; Sequence 38, Application US/09876348A						
; Patent No. US20020172951A1						
; GENERAL INFORMATION:						
; APPLICANT: Horwath, K. L. and Myers, K. L.						
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio						
; FILE REFERENCE: RB-125-R1						
; CURRENT APPLICATION NUMBER: US/09/876,348A						
; PRIOR FILING DATE: 2001-08-09						
; PRIOR FILING DATE: 2000-06-08						
; SOFTWARE: Microsoft Word						
; SEQ ID NO 38						
; LENGTH: 543						
; TYPE: DNA						
; ORGANISM: Tenebrio molitor						
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5						
US-09-876-348A-38						
Query Match 9.0%; Score 102; DB 10; Length 543;						
Best Local Similarity 88.1%; Pred. No. 3.8e-24;						
Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;						
QY	1	ATGGGCAGCAGCCCATCATCATCATCATGACGCGGCCCTGTGTCGCGCGGCGGCGGCCAT	60			
Db	64	ATGGGCAGCAGCCCATCATCATCATCATGACGCGGCCCTGTGTCGCGCGGCGGCGGCCAT	123			
QY	61	ATGGCTAGCATGACTGGTGGACAGCAATGGTGCGCGATCCAGCGCCTCTGTTTGTATG	120			
Db	124	ATGGTAGCATGACTGGTGGACAGCAATGGTGCGGATCCCTCACCGAACAGATA	183			

Qy 61 ATGGCTAGCATGACTGGTGGGACAGCAAAATGGGTTCGGGATCCAGGCGCCTCTGTTTGTATG 120
 |||||
 Db 124 ATGGCTAGCATGACTGGTGGGACAGCAAAATGGGTTCGGGATCCCTCACCAGCAACAGATA 183
 |||||
 Qy 121 GCTAAA 126
 |||||
 Db 184 CAGAAA 189

RESULT 13

```

US-09-876-796A-30
; Sequence 30, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifeedz Proteins and Method for Assaying Activity.
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876, 796A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 30
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; -OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4
US-09-876-796A-30

```

[illegible]

RESULT 14

```

US-09-876-796A-38
; Sequence 38, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ Antifreeze Proteins and Method for Assaying Activity.
; CURRENT APPLICATION NUMBER: US/09/876,796A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 38
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5
US-09-876-796A-38

```

```
Query Match          9.0%; Score 102; DB 10; Length 543;
Best Local Similarity 88.1%; Pred. No. 3.8e-24;
```

[illegible]

RESIT.T 15

```

US-10-228-063-43
; Sequence 43, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lananhan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Thermotoga neapolitana
US-10-228-063-43

```

	Query Match	9.0%;	Score 101.6;	DB 12;	Length 1436;
	Best Local Similarity	76.2%;	Pred. No. 9.7e-24;		
	Matches 125;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
QY	1	ATGGGCAGCAGCCATCATCATCATATAGCAGGGCTGGTGC	CGCGCGCAGCCAT	60	
Db					
Db	1	ATGGGCAGCAGCCATCATCATCATCAGCAGGGCTGGTGC	CGCGCGCAGCCAT	60	
QY	61	ATGGGTAGCATGACTGGTGGCAGCAAAATGGTCCGGATC	CAGCGCCTCTCTTTTGATG	120	
Db					
Db	61	ATGGGTAGCATGACTGGTGGCAGCAAAATGGTCCGGATC	CCCATGGCGAGTCTTCCCG	120	
QY	121	GCTAAACCCCAAGTGA	CTATCTTAGCCACAGGAGGCACCATCGC	164	
Db					
Db	121	GAGATCCCGAAGGTGCAGTTGCAGGGCAAGAGTCCCAACCT		164	

Search completed: August 27, 2003, 11:17:57
Job time : 847.661 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:35:29 ; Search time 3475.96 Seconds
(without alignments)
7922.115 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atgggcagcagccatcatca.....aagaaagggaattctttac 1133

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *

1: em_estba: *

2: em_esthum: *

3: em_estin: *

4: em_estmu: *

5: em_estov: *

6: em_estpl: *

7: em_estro: *

8: em_estl: *

9: gb_estl: *

10: gb_est2: *

11: gb_est3: *

12: gb_est4: *

13: gb_est5: *

14: gb_est6: *

15: em_estfun: *

16: em_estom: *

17: em_gss_hum: *

18: em_gss_inv: *

19: em_gss_pin: *

20: em_gss_vrt: *

21: em_gss_fun: *

22: em_gss_mam: *

23: em_gss_mus: *

24: em_gss_pro: *

25: em_gss_rod: *

26: em_gss_phg: *

27: em_gss_vrl: *

28: gb_gssl: *

29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	155.2	13.7	983	29	BZ568327
C 2	136	12.0	724	28	AQ990947
C 3	116.8	10.3	766	28	AQ500605
C 4	110.8	9.8	757	28	BH376509

C 5	108.6	9.6	843	28	AQ500657
C 6	106.8	9.4	768	29	BZ570599
C 7	97.8	8.6	553	28	AQ872693
C 8	97.8	8.6	553	28	AQ503225
C 9	96.2	8.5	596	28	AQ501032
C 10	96.2	8.5	630	28	AQ873860
C 11	94.4	8.3	453	28	AQ501722
C 12	94.4	8.3	459	28	AQ501723
C 13	93	8.2	561	28	AQ503362
C 14	89.6	7.9	663	28	BH371112
C 15	87	7.7	560	28	AQ873892
C 16	79.2	7.0	840	29	BZ571432
C 17	75.6	6.7	644	28	AQ517105
C 18	74	6.5	1378	28	BH770857
C 19	72.4	6.4	802	29	BZ571326
C 20	63.8	5.6	545	28	AQ873272
C 21	60.6	5.3	781	29	CNS06FWL
C 22	60.6	5.3	871	29	BZ571362
C 23	55	4.9	427	28	AQ991287
C 24	48	4.2	323	28	AZ049501
C 25	46.8	4.1	455	28	AZ049638
C 26	42.6	3.8	462	28	AQ502215
C 27	42.6	3.8	559	10	BG672925
C 28	41.8	3.7	388	12	BJ363543
C 29	41.8	3.7	492	12	BJ390828
C 30	41.2	3.6	260	12	BJ390253
C 31	41.2	3.6	405	12	BJ358893
C 32	41.2	3.6	473	12	BJ411665
C 33	41.2	3.6	516	12	BJ386769
C 34	41.2	3.6	554	12	BJ364097
C 35	41.2	3.6	554	12	BJ417890
C 36	41.2	3.6	559	12	BJ389977
C 37	41.2	3.6	574	12	BJ336311
C 38	41.2	3.6	588	12	BJ386962
C 39	41.2	3.6	594	12	BJ326765
C 40	41.2	3.6	605	12	BJ386822
C 41	41.2	3.6	619	12	BJ358849
C 42	41.2	3.6	622	12	BJ410498
C 43	41.2	3.6	623	12	BJ325097
C 44	41.2	3.6	623	12	BJ327615
C 45	41.2	3.6	626	12	BJ387400

ALIGNMENTS

RESULT 1
BZ568327/c
LOCUS
DEFINITION
pac82-164_7476.y2 pac82-164 Pseudomonas aeruginosa genomic clone
pac82-164_7476, genomic survey sequence.
ACCESSION
BZ568327
VERSION
BZ568327.1 GI:27201144
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 983)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol., (2002) in press
CONTACT: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

Location/Qualifiers

source
1. .983
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacc2-164_7476"
/clone_lib="pacc2-164"
/notes="clinical isolate 2-164 whole genomic shotgun library."
BASE COUNT 153 a 305 c 322 g 199 t 4 others
ORIGIN

Query Match 13.7%; Score 155.2; DB 29; Length 983;
Best Local Similarity 58.7%; Pred. No. 1.2e-34;
Matches 304; Conservative 0; Mismatches 210; Indels 4; Gaps 2;

QY 128 CCACAGTACTATCTCCTAGCCACAGGA-GGACCATCTGCTGCTGCGGGGAATCTAGGTC 186
DB 667 CCAACGTGTCTCTCTGCGCCACCGGCGGACCATCGCGGCGCGGCGCGCGCGCC 608
QY 187 AAGAG--TAGCTACTCTCTGAGCAGTCACCGTTGATAGCTTCTTTCGAGCGTCCCT 243
DB 607 AAGAGCGCCACCTACACCGCGCGCAAGGTGCGCGGTGACCGATTGCTGGCGAGGTACCG 548
QY 244 GCCATCAAGCAGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGCTCCCAAGAG 303
DB 547 CAGCTCAAGGACATCGCCACGTCGCGGCGGACAGGTGTTCCAGATCGCTCGGAAGC 488
QY 304 ATGACGGTGAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGCTCTTCGCCCAAAAA 363
DB 487 TTCACCAAGCAGAACCTGCTGGAACCTGGCAAGACCGTCGCCAAGCTGGCGCGATAGGCAC 428
QY 364 GAGACCGAAGCGGTGATCATCCCATGGAACCTGACACCATGGAAGAGCGCTTCTTC 423
DB 427 GAGTCGACGGCATCTGTGATCACCCAGGTACCGACACCTTGAAGAGACCGCTACTTC 368
QY 424 CTCACCTCACGGTGAAGGCAAAACCTGTCTGCGCTTGTAGCGCCATCGCTCCAGGC 483
DB 367 CTGACCTGTGTGAGCACACCGAGAGCGCTATCTGTGTGTGCTGATGCGCGCGGC 308
QY 484 TCTTCATGAGTCTGATGCGCCCATGAATCTCTATACCGCGTGAATGATGAGATCAAC 543
DB 307 ACCGCGATGTCCCGCAGCGCATGCTCAACCTGTACAAACCGCTGGCGGTGGCGGCGAC 248
QY 544 AAGACCTCTACTAACAAAGAGGTGATGATGATGAACCATGAGATTACGCGCGCAGA 603
DB 247 AAGTCGGCAGCGGCGAAGGCGGTGCTGATCACCATGAACGACGAGATCTCTCCGCGCG 188
QY 604 GAAGCGACCAAGCTCAACACCGCGCATGTCATCAATGCATT 641
DB 187 GACCGGACGAGATGCTCAACATCAAGACCGAAGCGTT 150

RESULT 2
AQ90947/c
LOCUS
DEFINITION
Photobacterium luminescens strain W14 M13 library
sequence.
AQ90947
AQ90947.1 GI:9649541
GSS.
Photobacterium luminescens
Photobacterium luminescens
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photobacterium.
REFERENCE
1 (bases 1 to 724)
french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,
Daborn, P.J., Bowen, D. and Blattner, F.R.
A genomic sample sequence of the entomopathogenic bacterium
Photobacterium luminescens W14: potential implications for virulence
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
20378633

PUBMED
COMMENT

10919786
Contact: french-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bsr@c Bath.ac.uk

This is one of 2,122 random reads from the M13 library. For
annotation of identified clones (BLASTX, BLASTN and mapping to E.
coli K12 genome) please see french-Constant et al. 2000, Nucleic
Acids Res.

Seq primer: M13 Forward
Class: shotgun.

FEATURES
source

Location/Qualifiers
1. .724

/organism="Photobacterium luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/clone="PLG01794"
/dev_stage="primary phase variant"
/clone_lib="Photobacterium luminescens strain W14 M13
library"

BASE COUNT 218 a 176 c 137 g 192 t 1 others
ORIGIN

Query Match 12.0%; Score 136; DB 28; Length 724;
Best Local Similarity 60.8%; Pred. No. 5.3e-29;
Matches 242; Conservative 0; Mismatches 150; Indels 6; Gaps 1;

QY 118 ATGCTAAACCCCAAGTGACTATCTAGCCACAGAGGACCATCGCTGTTTCGGGGAA 177
DB 404 ATTGCTTTACCAATATTACACTTTTGGCGACGGGGGAGTAATTGCGGTGGCGGAAG 345
QY 178 TCTAGCGTCAAGAGTAGCTACTCTGCTGAGCAGTCACCGTTGATAAGCTTCTTGACGCC 237
DB 344 TCTGCAATCCATCCATTAATGTTGCTGGAAGGTTGCTATTGATGAATTAGTCAATGCT 285
QY 238 GTCCCTGCAATCAACGACCTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGCTCC 297
DB 284 GTTCTGAGATCAAAAATATCGCAACCTTAAAGGTGAGCAGGTAGTTAGTATTGTTGCG 225
QY 298 CAAGAGATGAGCGGTAAAGTGTGCTTAAATAGCCAAAGCGTGTCAATGAGCTCTCTCGCC 357
DB 224 CAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 165
QY 358 CAAAAGAGACCGGAAGCGGTGATCATCACCATGGAACCTGACACCATGGAAGACCGCT 417
DB 164 AAA-----ACCGATGGTTTGTATACCCACCGTACTGACACCATGGAAGACCGCT 111
QY 418 TTCTTCCTCAACCTCAGCGTGAAGCAAAACCTGTCTGCTTGTAGCGCGCATGCGT 477
DB 110 TACTTCTCGATCTCACTACCACTCCAGAACCGGATGATGATGATGATGATGATGATGATGAT 51
QY 478 CCAGGCTTTCATGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 515
DB 50 CCATCTACGGCGCTTGGTGGCGATGGCGCGTAAATCT 13

RESULT 3
AQ500605/cLOCUS
DEFINITION

AQ500605
V36C6 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
genomic 5', genomic survey sequence.

ACCESSION
VERSIONKEYWORDS
SOURCE

ORGANISM

766 bp DNA linear GSS 29-APR-1999
Photobacterium luminescens strain W14 M13 library
sequence.
AQ500605
AQ500605.1 GI:4706315
GSS.
Saccharomyces cerevisiae
Saccharomyces cerevisiae (baker's yeast)
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;


```

QY 324 TAAACTAGCCCAAGCGTGTCAATGAGCTCTCGCCCAAAAGAGACCGGCGGTGATCAT 383
Db 523 CCAGCTCACTGCTGCTGTGCAAGCGGTATCAACAGTCAGACGTTCTGGCGGTGATCGT 464
QY 384 CACCCATGGAACCTGACACCAATGGAAGAGACCGCTTTCTCTCAACCTCAACGCTGAAAG 443
Db 463 TTCTCATGGCACAGATACGCTGGAGGAACAGCTTTTGGCTTGATCTGACATGAAGTC 404
QY 444 CCAAAACCTGTCTGCGCTTGTAGCGCCATGCGTCCAGGCTCTTCATGAGTCTGATGG 503
Db 403 CATAAACCAATCGTCTGATTTGGCGCACAGCGCAACGCGCTCTTTCAGACTTTGACGG 344
QY 504 CCCATGAATCTCTATAACCGCGTGAATGTAGCGATCAACAAAGCCTCTACTAACAAAG 563
Db 343 TCCGCGCAACTTGTCTTAACGCTGTGCGCATCGTGTGTATCCCAATCCAGATCGTGG 284
QY 564 AGTGGTGATGTGATCAACGATGAGATTACGCGCGCCAGAGAGACCAAGCTCAACAC 623
Db 283 CGCATGCTGGCGATGAACCAACGATCAACGCTACCAAGTATGTCAACAGACACAC 224
QY 624 CACCGCAGTCAATGCTTTGCTTTCGCGCCCAACACAGGTAAATCGGCCTTATTTGGTGG 683
Db 223 TGCACACGTTGAACCTTTTAAGTCTGGTGTGCTGCTTATTTGGTGGTCTACCCAGA 164
QY 684 CAAAGTCGAGTATTTCACTCAATTCGCTTCGACCTCACACCTTGAAGTGTGAT 741
Db 163 CAAGGTCTTTACCGCACTAGCGCGATCGCAATTTAACTTTCTATCAAGGCAGAT 106

```

RESULT 5

AQ500657/c

LOCUS

AQ500657 843 bp DNA linear GSS 29-APR-1999
 V34rII mTn-3XHA/lacZ Insertion Library Saccharomyces cerevisiae
 genomic 5', genomic survey sequence.

ACCESSION

AQ500657

VERSION

AQ500657.1

KEYWORDS

GI:4706367

SOURCE

GSS.

ORGANISM

Saccharomyces cerevisiae (baker's yeast)

REFERENCE

Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

AUTHORS

1 (bases 1 to 843)
 Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
 deStagis, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
 Unenky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

TITLE

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption

JOURNAL

Unpublished

COMMENT

Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3XHA/lacZ insertion.
 Seq primer: GGCCTTCTTTGGAGTAC
 Class: transposon-tagged.

FEATURES

source

1. .843
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3XHA/lacZ Insertion Library"
 /note="Vector: pHS6-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHS6-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mTn-3XHA/lacZ
 minitransposon containing lacZ, URA3, and tet resistance."
 245 a 182 c 182 g 224 t 10 others

ORIGIN

Query Match 9.6%; Score 108.6; DB 28; Length 843;
 Best Local Similarity 50.7%; Pred. No. 8.6e-21;
 Matches 255; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

```

QY 208 GCAGTCACCGTTGATAAGCTTCTTGACGCGTCCCTCGCCATCAACGACCTAGCCACCATC 267
Db 679 GGATTAACCGTAATGATNTAATAGAGCGGTCATCTTTAGCTGAGAGGACATCTG 620
QY 268 AAGGTGAACAGATCTCAAGCATTTGCTCCCAAGAGATACGGGTAAAGTGTGGCTTAA 327
Db 619 GACTATCTTCAAGTGTCTAACGTTGGTTTCANATTCTTTANACTATATACGATCTGATCC 560
QY 328 CTAGCCCAAGCGTGTCAATGAGCTCTCGCCCAAAAGAGACCGAAGCCCTGATCATACC 387
Db 559 TTGTATCAGGTTATCTCCGAGCATTAGCTTCTGTAGTACTAGCTGGTGGTGTGCACT 500
QY 388 CATGGAATGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCACGGTGAAGGCCAA 447
Db 499 CATGNGACCGACCATATGAGAGAGACAGCTTTCTTCTTAGATTGACCATATAATTGAG 440
QY 448 AAACCTGTCTGCTTGTAGCGCCCATGCTCCAGGCTCTTCCATGAGTGTGATGGCCCC 507
Db 439 AAGCCAGTATGTATCGCAGCGCTATGGTCCGACCTGCGCACGCTCTGCTGATGGCCCC 380
QY 508 ATGAATCTCTATAACCGCGTGAATGTAGCGATCAACAAAGCCTCTACTAACAAAGAGTG 567
Db 379 ATGAATTTATATCAAGCAGTGTCTATTGCTGCTTCTGAGAATCACTGGGTGCTGGCAG 320
QY 568 GTGATTGTGATGAACGATGAGATTCAAGCCGCCAGAGAGAGCGCAACAGCTCAACACCA 627
Db 319 ATGACTACTCTAAACGATCGTATTTCCTCTGGGTTTGGACACGAAATGAATGCCAAC 260
QY 628 GCAGTCAATGATTTGCTTTCGCCCAACACACAGGTAAATCGGCACAGTCTATTATGGCAA 687
Db 259 TCTTTAGATACATTACAGCGGATGAACAGGATATTAGGTACTTTTCAATATGATGAC 200
QY 688 GTCAGTATTCTCACTCAATCCGT 710
Db 199 GTGAGGTTTACTACCCACCAGT 177

```

RESULT 6

BZ570599/c

LOCUS

BZ570599 768 bp DNA linear GSS 17-DEC-2002
 msh2_1446.y2 msh Pseudomonas aeruginosa genomic clone msh2_1446,
 genomic survey sequence.

ACCESSION

BZ570599

VERSION

BZ570599.1

KEYWORDS

GI:27205660

SOURCE

GSS.

ORGANISM

Pseudomonas aeruginosa

REFERENCE

Pseudomonas aeruginosa

AUTHORS

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

TITLE

1 (bases 1 to 768)

JOURNAL

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
 Burns, J.L., Kaul, R. and Olsen, M.V.

COMMENT

Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

FEATURES

source

1. .768
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"

/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_1446"
/note="Environmental isolate. Whole genomic shotgun library."
library."

BASE COUNT 116 a 246 c 239 g 165 t 2 others
ORIGIN
Query Match 9.4%; Score 106.8; DB 29; Length 768;
Best Local Similarity 53.6%; Pred. No. 2.8e-20;
Matches 287; Conservative 0; Mismatches 243; Indels 5; Gaps 3;

QY 331 GCCAAGCGTCAATGAGCTCTCGCCCAAAAGAGACCGAGCGTGTATCATCCCAT 390
Db 610 GCAATACCGTGCCTAGCTGCGGATACGAAACAGCTCGCGGATCATCCCAT 551
QY 391 GGAATCGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCACGCTGAAAGCCAA 450
Db 550 GGTACCGACACCT-GAAGAGACCGCCCACTTCTGACCTGTGTGAGCACACCGGAGAG 492
QY 451 CTTGTCTGCTGTAGGCGCATGCTCCAGCTCTTCCATGAGTGTCTGATGGCCCATG 510
Db 491 CCTATCGTGTGTGCTGCTGATGCGCCCGGACCGCNATGTCCGCGCATGCTC 432
QY 511 AATCTCTAATACCGCTGATGTAGCGATCAACAAAGCTCTACTAACAAGAGAGTGTG 570
Db 431 AACCTGTACAACCGC-GCGGTGCGCGGACAGTCCGACCGCGGCAAGGCGTGTG 373
QY 571 ATTGTGATGAACGATGAGATTACGCGCGGACGAGAGCGACCAAGCTCAACACCGCA 630
Db 372 ATCACCATGAACGACGAGATCTCTCCGCGCGGACGCGCAAGATGTCAACATCAAG 313
QY 631 GTCATGATTTGCTTCCGCGCAACAGAGTAAATCGGCACAGTCTATTATGGCAAGTC 590
Db 312 ACCGAGCGCTT---CAAGAGCGCTGGGGCCGCTGGCGATGCTGTGAGGCGCAAGAGC 256
QY 691 GAGTATTTCACTCAATCCGTTGACCTCACCCCTTGCAAGTGTGATTTAGTATTA 750
Db 255 TACTGTGTCGGCACCGGTGAAGCGGACACGGTCACTCCGATCCGATTCACATCAAGCAG 196
QY 751 ATCGAAGAACTCCCGAGTGCATATTTTACGCTCACCCGATGATGATGATTTT 810
Db 195 ATCTCGCGCTGCTCGGTGGAATCGCTACAGCTACGGAACGCTCAGCGACACCGCC 136
QY 811 GTCATGAGCCCTTCCGCGAGAGCAAGGATCATCATGCGCATGCGCATGGCA 865
Db 135 TACAAGGCGCTGCGCCAGCGCGCGCAAGGCGATCATCATGCGCGAGGGGA 81

RESULT 7
AQ872693/c
LOCUS V34E5 mTn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces
DEFINITION cerevisiae genomic 5', genomic survey sequence.

ACCESSION AQ872693
VERSION GSS.
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 553)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL Unpublished
CONTACT: Kumar A
COMMENT Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTCTTTCTTTGGAAGTAC
Class: transposon-tagged.
Location/Qualifiers

FEATURES
source

1..553
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="AB972 - trp1 r(0) (S288C background)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"
/note="Vector: pHS6-Sal; A yeast genomic DNA library was
prepared in pHS6-Sal; genomic DNA was size-fractionated
(DNA of roughly 2-3 kb in length) prior to cloning. This
library was subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 152 a 127 c 124 g 150 t
ORIGIN

Query Match 8.6%; Score 97.8; DB 28; Length 553;
Best Local Similarity 53.9%; Pred. No. 1.1e-17;
Matches 201; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 338 GTGTCAATGAGCTCTCGCCCAAAAGAGACCGAGCGTGTATCATCCCATGGAAC 397
Db 548 GTATCTCCGAGCACCTAGCTTCTGATGACTACGCTGTGCTGCTGCTGCTGCTG 489
QY 398 ACACCATGAAGAGACCGCTTTCTTCTCAACTCACGCTGAAAGCCCAAAACCTGTCT 457
Db 488 ACATATGAGAGAGACAGCTTTCTTCTAGATTGACCAATTAATTCAGAGACCGAT 429
QY 458 GCCTTTGAGGCGCCATGCTCCAGGCTTTCCATGAGTGTCTGATGCCCATGAATCTCT 517
Db 428 GTATCGAGCGCTATGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
QY 518 ATAAAGCGCTGAATGTAGCGATCAACAAGCTCTTACTAACAAGAGTGTGATTTGA 577
Db 368 ATCAAGCAGTGTATTGTCTGCTGAGAAATCACTGGGTGCTGCGACGATGATCACTC 309
QY 578 TGAACGATGAGATTACGCGCGCAGAGAGACCAAGCTCAACACCGCAGTCAATG 637
Db 308 TAAACGATCGTATTGCTCTGGGTTTGGACAAAGAAATGAATGATCCAACTCTTAGATA 249
QY 638 CATTTGCTTTCGCCCAACACAGGTAAATCGGCACAGTCTATTATGCGAAGTCGATTT 697
Db 248 CATTCAGAGCGGATGAACAGGGATATTAGGTATTCTTTTCAATGATGACGTGGAGTTT 189
QY 698 TCACTCAATCCGT 710
Db 188 ACTACCCACCACT 176

RESULT 8
AQ503225/c
LOCUS V34E05 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
DEFINITION genomic 5', genomic survey sequence.

ACCESSION AQ503225
VERSION GSS.
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 553)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,

TITLE Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
 JOURNAL Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 COMMENT Gene Disruption
 Unpublished
 Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of min-3xHA/lacZ insertion.
 Seq primer: GGCCTCTCTTTTGAAGTAC
 Class: transposon-tagged.
 Location/Qualifiers

FEATURES
 source
 1..553
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3xHA/lacZ Insertion Library"
 /note="Vector: pHSS6-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHSS6-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mTn-3xHA/lacZ
 minitransposon containing lacZ, URA3, and tet resistance."
 BASE COUNT 152 a 127 c 124 g 150 t
 ORIGIN

Query Match 8.6%; Score 97.8; DB 28; Length 553;
 Best Local Similarity 53.9%; Pred. No. 1.1e-17;
 Matches 201; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
 QY 338 GTGTCATAGCTCTCGCCCAAAAGAGACCGGTGATCATCACTGGAATG 397
 Db 548 GTATCTCCGAGGCACTAGCTTCTGATGACTACGCTGTGGGTGTCTACTATGGACG 489
 QY 398 ACACCATGGAGAGACCGCTTCTCTCAACCTCAGCGTGAAGAACCAAAACCTGCT 457
 Db 488 ACATATGAGGAGACACTTCTCTTAGATTGACATTAATTCAGAGGCAATAT 429
 QY 458 GCCTTGTAGCGGCATCGCTCCAGGCTCTTCCATGAGTGTGATGGCCCCATGAATCTCT 517
 Db 428 GTATCGCAGCGCTATCGCTCCAGCACTGCACGCTCTGCTGATGGCCCAATGAATTAT 369
 QY 518 ATAAAGCGCGAATGATGCGCATCAACAGCCTCTACTACAAAGGAGTGTGATTGTA 577
 Db 368 ATCAAGCAGTGTCTATTGCTGCTTCTGAGAAATCACTGGGTCGTCGACGATGATCACTC 309
 QY 578 TGAACGATGAGATTACGCGCCAGAGAGCGACCAAGCTCAACACCAACCGCAGTCAATG 637
 Db 308 TAAACGATCGTATTGCTCTGGGTTTGGACAGGAAATGAATGCCAATCTTTAGATA 249
 QY 638 CATTTGTTGCGCCACACAGGTAATAATCGGCACAGCTCTATTATGGCAAAAGTCGAGTATT 697
 Db 248 CATTCAGCGGATGAACAGGGATATTAGGTACTTTTCAAAATGATGACGTGGAGTTT 189
 QY 698 TCACTCAATCCGT 710
 Db 188 ACTACCACCACT 176

RESULT 9
 AQ501032/c
 LOCUS V28H11 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
 DEFINITION genomic 5', genomic survey sequence.
 ACCESSION AQ501032
 VERSION AQ501032.1
 KEYWORDS GI:4706682
 GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae

REFERENCE
 AUTHORS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 1 (bases 1 to 596)
 Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
 desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption
 Unpublished
 Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3xHA/lacZ insertion.
 Seq primer: GGCCTCTCTTTTGAAGTAC
 Class: transposon-tagged.
 Location/Qualifiers

FEATURES
 source
 1..596
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3xHA/lacZ Insertion Library"
 /note="Vector: pHSS6-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHSS6-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mTn-3xHA/lacZ
 minitransposon containing lacZ, URA3, and tet resistance."
 BASE COUNT 171 a 132 c 130 g 163 t
 ORIGIN

Query Match 8.5%; Score 96.2; DB 28; Length 596;
 Best Local Similarity 54.0%; Pred. No. 3.4e-17;
 Matches 197; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
 QY 346 GAGCTCTCGCCCAAAAGAGACCGGTGATCATCACTGGAATGGAACATG 405
 Db 533 GAGGCATAGCTTCTCTGATGATTACGCTGTGGGTGTCTACTATGGACCGACATG 474
 QY 406 GAAGAGACCGCTTCTCTCAACCTCAGCGTGAAGAACCAAAACCTCTCTGCTTGTGA 465
 Db 473 GAGGAGACGCTTCTCTTAGATTGACATAAATTCAGAGACCGATGATGATCGCA 414
 QY 466 GCGCGCATGCTCCAGGCTCTTCCATGATGCTGATGGCCCATGAATCTCTATAACGCC 525
 Db 413 GCGGTATGCTCCAGCCACTGCCACGCTGCTGATGGCCCAATGAATTTATATCAAGCA 354
 QY 526 GTGAATGTAGCGATCAAAAGCCTCTACTAAAGAGAGTGTGTGATGAACGAT 585
 Db 353 GTGTCTATTGTGCTTCTTGAGAAATCACTGGGTGCTGGCAGCATGATCACTCTAAACGAT 294
 QY 586 GAGATTTCACGCGCCAGAGAGACCGACAGCTCAACACCCAGCAGTCAATGATTTGCT 645
 Db 293 CGTATGCTCTGGGTTTGGACAAAGAAATGAATGCCAATCTTTTAGATACATTCAGA 234
 QY 646 TCGCCCCACACAGGTAAATCGGCACAGTCTATTATGGCAAAAGTCAGATTTTCACTCAA 705
 Db 233 CGGGATGAACGGGATATTAGTTACTTTTCAATGATGACGTGGAGTTTACTACCCA 174
 QY 706 TCCGT 710
 Db 173 CCAGT 169

RESULT 10
 AQ873860/c
 LOCUS V82E1 mTn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces
 DEFINITION

cerevisiae genomic 5', genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AQ873860
AQ873860.1 GI:6286104

GSS.
Saccharomyces cerevisiae (baker's yeast)

Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 630)
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption

JOURNAL
COMMENT

Unpublished
Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.

Seq primer: GGCCTCTCTTCTTGGAGTAC

Class: transposon-tagged.

FEATURES

source

1..630
/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="AB972 - trp1 r(0) (S288C background)"

/db_xref="taxon:4932"

/lab_host="E. coli"

/clone_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"
/note="Vector: pHS6-Sal; A yeast genomic DNA library was
prepared in pHS6-Sal; genomic DNA was size-fractionated
(DNA of roughly 2-3 kb in length) prior to cloning. This
library was subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT
ORIGIN

179 a 138 c 138 g 175 t

Query Match 8.5%; Score 96.2; DB 28; Length 630;

Best Local Similarity 53.6%; Pred. No. 3.5e-17;

Matches 200; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 338 GTGTCAATGAGTCTCGCCAAAAGAGACCGAGCGGTGATCATCACCCTGGAACGTG 397

DB 550 GTATCTCCGAGGCACTAGCTTCTGTGATGACTACGCTGGTGTGTTCACTCATGCGACCG 491

QY 398 ACACCATGAGAGACCGCTTCTCTCACTCAGGTGAAAGCCAAACCTGTCT 457

DB 490 ACACATGAGGAGACAGCTATCTCTTAGATTGACCAATAATTCAGAGAAGCCAGTAT 431

QY 458 GCCTTTGAGCGGCACTCGCTCCAGGCTCTTCCATGAGTGTCTGATGCCCCCATGAATCTCT 517

DB 430 GTATCCAGCGGCTATGCGTCCAGCACTGCCACGCTGCTGATGCCCAATGATTTAT 371

QY 518 ATAAACCCGTAATGATGAGGATCAACAAAGCCCTCTACTAAAGAGATGGTATTGTGA 577

DB 370 ATCAAGCAGTGTCTATTGCTGTCTCTGAGAAATCACTGGGTCTGTCGACGATGATCACTC 311

QY 578 TGAACCATGAGTATCAAGCGCCAGAGAGAGCGACCAAGCTCAACACCGAGTCAATG 637

DB 310 TAAACCATGATTTGCTCTGGGTTTGGACAACGAAATGAATGCCAATCTTTAGATA 251

QY 638 CATTTGCTCGCCCAACACAGGTGAAATCGGCACAGTCTATTATGGCAAGTGCAGTATT 697

DB 250 CATTCAGCGGATGACAGGGATATTAGTTACTTTTCAATGATGACGTGAGTTT 191

QY 698 TCACCTCAATCCGT 710

Db 190 ACTACCCACCAGT 178

RESULT 11

AQ501722/c

LOCUS

DEFINITION

VS15G1 mTn-3xHA/lacZ Insertion Library Saccharomycetes cerevisiae

genomic 5', genomic survey sequence.

ACCESSION

AQ501722

VERSION

AQ501722.1 GI:4707372

KEYWORDS

Saccharomyces cerevisiae (baker's yeast)

SOURCE

Saccharomycetes cerevisiae

ORGANISM

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE

1 (bases 1 to 453)

AUTHORS

Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,

desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,

Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,

Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and

Gene Disruption

JOURNAL

COMMENT

Unpublished

Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.

Seq primer: GGCCTCTCTTCTTGGAGTAC

Class: transposon-tagged.

FEATURES

source

1..453

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

/lab_host="E. coli"

/clone_lib="mTn-3xHA/lacZ Insertion Library"

/note="Vector: pHS6-Sal; A yeast genomic DNA library

(lacking mitochondrial DNA) was prepared in pHS6-Sal;

genomic DNA was size-fractionated (DNA of roughly 2-3 kb

in length) prior to cloning. This library was

subsequently mutagenized with a mTn-3xHA/lacZ

minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT

111 a 121 c 115 g 105 t

ORIGIN

Query Match 8.3%; Score 94.4; DB 28; Length 453;

Best Local Similarity 56.9%; Pred. No. 9.8e-17;

Matches 173; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 338 GTGTCAATGAGTCTCGCCAAAAGAGACCGAGCGGTGATCATCACCCTGGAACGTG 397

DB 372 GTATCTCCGAGGCACTAGCTTCTGTGATGACTACGCTGGTGTGTTCACTCATGCGACCG 313

QY 398 ACACCATGGAAGAGACCGCTTCTCTCAACCTCACGGTGAAGCCAAACCTGTCT 457

DB 312 ACACATGAGGAGACAGCTTCTCTTAGATTGACCAATAATTCAGAGAAGCCAGTAT 253

QY 458 GCCTTTGAGCGCCATGCGTCCAGGCTTTCATAGTGTGATGCCCCCATGAATCTCT 517

DB 252 GTATCGAGGCGTATGCGTCCAGCAGCTGCCACGCTGCTGATGGCCCAATGAATTTAT 193

QY 518 ATAAACCCGTAATGATGAGGATCAACAAAGCCCTCTACTAAAGAGATGGTATTGTGA 577

DB 192 ATCAAGCAGTGTCTATTGCTGTCTCTGAGAAATCACTGGGTCTGTCGACGATGATCACTC 133

QY 578 TGAACCATGAGTATCAAGCGCCAGAGAGAGCGACCAAGCTCAACACCGAGTCAATG 637

DB 132 TAAACCATGATTTGCTCTGGGTTTGGACAACGAAATGAATGCCAATCTTTAGATA 73

```

QY      638 CATT 641
Db      |||||
        72 CATT 69

RESULT 12
AQ501723/c
LOCUS   459 bp DNA linear GSS 29-APR-1999
DEFINITION V15G2 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
           genomic 5', genomic survey sequence.
ACCESSION AQ501723
VERSION   AQ501723.1 GI:4707373
KEYWORDS GSS.
SOURCE   Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
1 (bases 1 to 459)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL Unpublished
COMMENT Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTCTCTTCTTGAAGTAC
Class: transposon-tagged.
FEATURES
source
1..459
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
(lacking mitochondrial DNA) was prepared in pHS6-Sal;
genomic DNA was size-fractionated (DNA of roughly 2-3 kb
in length) prior to cloning. This library was
subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT 116 a 119 c 120 g 102 t
ORIGIN

Query Match 8.3%; Score 94.4; DB 28; Length 459;
Best Local Similarity 56.9%; Pred. No. 9.9e-17;
Matches 173; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY      338 GTCTCATGAGCTCTCTGCCCAAAAGACGCGGTGATCATCAACCCATGAACTG 397
Db      |||||
        349 GTATCTCCGAGGCATGACCTCTGATGACTACGCTGTGTGCGGTGTCTACTCATGGACCG 290

QY      398 ACACCATGGAAGAGACCGCTTCTCTCAACCTCAGGTGAAAGCCAAACCTGTCT 457
Db      |||||
        289 ACATATGAGGAGACAGCTTCTTCTTAGATTGACCATTAATTCAGAGAAGCCGAT 230

QY      458 GCCTTTGAGCGCCATCGCTCCAGGCTCTTCCATGAGTGTGTGATGGCCCAACCTGTCT 517
Db      |||||
        229 GTATCGAGCGGTATCGCTCCAGCCACTGCCAGCTCTGCTGATGGCCCAATGATTAT 170

QY      518 ATAACCCGCTGATGTAGGATCAACAAAGCTCTACTAACAAGAGAGTGTGTGATGGA 577
Db      |||||
        169 ATCAAGCAGGTCTATTGTCTTCTGAGAAATCACTGGGTGCGGACCATGATCACTC 110

QY      578 TGAACGATGAGATTACGCGGCCAGAGAGACCAAGCTCAACACACCGCAGCTCAATG 637
Db      |||||

```

```

Db      109 TAAACGATCGTATTGCTCTGGGTTTTTGGACAGAAATGAATGCCAATCTTTAGATA 50
QY      638 CATT 641
Db      |||||
        49 CATT 46

RESULT 13
AQ503362/c
LOCUS   561 bp DNA linear GSS 29-APR-1999
DEFINITION V57D10 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
           genomic 5', genomic survey sequence.
ACCESSION AQ503362
VERSION   AQ503362.1 GI:4705908
KEYWORDS GSS.
SOURCE   Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
1 (bases 1 to 561)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL Unpublished
COMMENT Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTCTCTTCTTGAAGTAC
Class: transposon-tagged.
FEATURES
source
1..561
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
(lacking mitochondrial DNA) was prepared in pHS6-Sal;
genomic DNA was size-fractionated (DNA of roughly 2-3 kb
in length) prior to cloning. This library was
subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT 158 a 128 c 118 g 157 t
ORIGIN

Query Match 8.2%; Score 93; DB 28; Length 561;
Best Local Similarity 53.4%; Pred. No. 3e-16;
Matches 195; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY      346 GAGCTCTCGCCAAAAGAGACCGGCGGTGATCATCAACCCATGAGACTGACACCATG 405
Db      |||||
        514 GAGGATTAGCATATGATGATTACGCTGTGTGCGGTGTCTACTCATGGACCGACATATG 455

QY      406 GAAGAGACCGCTTCTTCTCAACCTCAGGTGAAAGCCAAACCTGTCTGCTTGTGA 465
Db      |||||
        454 GAGGAGACAGCTTCTTCTTAGATTGACCATAGATTACAGAGACCATGATGATCGCA 395

QY      466 GCGCCCATGCGTCCAGGCTCTTCCATGAGTGTGTGATGGCCCAATGATCTCTATACGCC 525
Db      |||||
        394 GCGCTATGCGTCCAGCCACTGCGCTGTGTGATGGCCCAATGATTTATATCAAGCA 335

QY      526 GTGAATGTAGCGATCAACAAAGCCCTCTACTAACAAGAGAGTGTGTGATGTAACGAT 585
Db      |||||
        334 GTGTCTATTGATGCTTCTTGAGAAATCACTGGGTGCGGACCATGATGATCACTTAAACGAT 275

```


QY 586 GAGATTACGCGCCGAGAGAGCGCAAGCTCAACACCGCGAGTCAATGCAATTGCT 645
 |||||
 Db 274 CGTATTGCTCTGGGTTTGGACACGAAATGAATGCCAACTCTTTAGATACATTAGA 215
 |||||
 QY 646 TCGCCCAACACAGGTAAATCGGCACACTCTATTATGCGAAAGTGCAGTATTCTACTAA 705
 |||||
 Db 214 GCGATGACAGGGATATTAGGTACTTTTCAATGATGACGCGAGGTTTACTACCCA 155
 |||||
 QY 706 TCCGT 710
 |||||
 Db 154 CCACT 150
 |||||

RESULT 14
 BH371112/c
 LOCUS
 DEFINITION AG-ND-137P15.TF ND-TAM Anopheles gambiae genomic clone AG-ND-137P15
 , genomic survey sequence.

ACCESSION BH371112 GI:17317237
 VERSION BH371112.1
 KEYWORDS GSS.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 Anopheles.

REFERENCE 1 (bases 1 to 663)
 AUTHORS Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren
 ,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B., Gardner,M.J.
 and Collins,F.H.

TITLE Construction of a BAC library and generation of BAC end
 sequence-tagged connectors for genome sequencing of the African
 malaria mosquito Anopheles gambiae

JOURNAL Mol. Genet. Genomics 268 (6), 720-728 (2003)

MEDLINE 22542063

PUBMED 1265398

COMMENT Other_GSSs: AG-ND-137P15.TR

Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@tigr.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by
 F.H. Collins and sequenced by The Institute for Genomic Research
 (TIGR). The BAC library was generated from A. gambiae PEST strain
 DNA. All DNA was extracted from newly hatched first instar larvae
 to minimize the inclusion of DNA from microorganisms that inhabit
 the gut. The DNA is derived from mixed sexes of larvae. The BAC
 library was constructed at Texas A&M University BAC Center
 University, College Station, Texas 77843-2123, USA using a HindIII
 partial digest.

Seq primer: M13 For

Class: BAC ends.

Location/Qualifiers

1. .663
 /organism="Anopheles gambiae"
 /mol_type="genomic DNA"
 /strain="PEST"
 /db_xref="taxon:7165"
 /clone_lib="AG-ND-137P15"
 /notes="Vector: pECBAC1; Site 1: HindIII"
 160 a 175 c 168 g 160 t

BASE COUNT 160 a 175 c 168 g 160 t

ORIGIN

Query Match 7.9%; Score 89.6; DB 28; Length 663;
 Best Local Similarity 50.4%; Pred. No. 3.4e-15;
 Matches 244; Conservative 0; Mismatches 239; Indels 1; Gaps 1;

QY 259 GCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGTAAGTG 318
 |||||

Db 589 GCCAAGATCCAGGTAAACAACTTATCCACGTTCCATCGGACTACATGGGCGCTGAGCGC 530
 |||||
 QY 319 TGGCTTAACTAGCCAGCGGTGTCATGAGCTCTCGCCCAAAAGAGACCGAAGCCGTG 378
 |||||
 Db 529 TGGGTCCAGTCTACTGCTGTGTCGAAGCGCTATCAACAAGTCAGACGTTGCTGGCGTG 470
 |||||
 QY 379 ATCATCACCACATGGAATGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCACCGTG 438
 |||||
 Db 469 ATCGTTTCTCATGGCACAGATACGCTGGAGAAACAGCTTTTGGCTGATCTGACAGTG 410
 |||||
 QY 439 AAAAGCCAAAACCTGCTGCTGCTAGGCGCCATGCGTCCAGGCTCTTCCATGAGTGCT 498
 |||||
 Db 409 AAGTCCAATAAACCAATCGTCTGATTGGCGCAGCGCAACGCTCGGTTTCAGACTTT 350
 |||||
 QY 499 GATGGCCCATGAATCTCTATACGCGGTGAATGTAGCGATCAACAAGCGCTCTACTAAC 558
 |||||
 Db 349 GACGGTCCCGGCAACTTGTCTTAACGCTGTGGCATGCTGTGATCCCAATCCCAAGGAT 290
 |||||
 QY 559 AAAGGAGTGGTGTGATGTGATGAACGATGAGATTTCACGCCGCCAGAGAGCGCAAGCTC 618
 |||||
 Db 289 CGTGGGCCCATGCTGGCGATGAACACCAAGATCAACGCTACCAAGTATGTACCAGACA 230
 |||||
 QY 619 AACACACCGCAGTCAATGCAATTT-GCTTGGCCCAACACAGGTAAATTCGACAGTCTA 677
 |||||
 Db 229 CACACTGCCCAAGCTTGAACACTTTTAAAGTTCCTGGTGAAGTTCGGCTTTTATGTTGAGGTTCTA 170
 |||||
 QY 678 TTATGCAAGTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCCCTTGAAGTGAAGTT 737
 |||||
 Db 169 CCCAGACAAAGTCGTTTACCACACTAGCGCATGCCCAATTTAAACTTTCTTATCAAGGC 110
 |||||
 QY 738 TGAT 741
 |||||
 Db 109 AGAT 106
 |||||

RESULT 15

AQ873892/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE.

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .560

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="AB972 - trp1 r(0) (S288C background)"

/db_xref="taxon:4932"

/lab_host="E. coli"

/clone_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"

Seq primer: GGCCTTCTTCTTTTGGAGATAC

Class: transposon-tagged.

Location/Qualifiers

1. .560

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="AB972 - trp1 r(0) (S288C background)"

/db_xref="taxon:4932"

/lab_host="E. coli"

/clone_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"

/note="Vector: pHS6-Sal; A yeast genomic DNA library was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT		155 a	127 c	119 g	158 t	1 others
ORIGIN						
Query Match						
Best Local Similarity 7.7%; Score 87; DB 28; Length 560;						
Matches 177; Conservative 0; Mismatches 151; Indels 0; Gaps 0;						
QY	383	TCACCCATGGA	CTGACACCATG	GAAGAGACCG	CTTCTTCCTCA	ACCTCACGGTGA
DB	503	TCACCTCATGG	ACCGACACTAT	GGAGAGACGCT	ATTTCTTAGAT	TTGACCATANAT
QY	443	GCCAAAACTGT	CTGTGCTTGT	AGGCGCCATG	CGTCCAGGCTC	TTCCATGAGTGT
DB	443	CAGAGAGCG	AGTATGATTC	GCAGGGGATAT	GGTCCAGCCACT	GGCCACTGCTGT
QY	503	GCCCCATGA	ATCTCTATA	AGCCGTGAAT	GTAGCGATCAA	CAAGCCTCTACT
DB	383	GCCCAATGA	ATTTATATCA	AGCAGTGTCT	ATTGTCTTCTG	AGAAATCACTGG
QY	563	GAGTGTGAT	GTGATGAAC	GATCAGATTCA	CGCCGCCAGAG	AGCGACCAAGCT
DB	323	GCACGATGA	TCACTCTAA	ACGATCGTAT	TGCCCTCTGGG	TTTGGACAACGAA
QY	623	CCACCGCAG	TCAATGTCT	TGCCCCAAC	ACAGGTAAAT	CGGCACAGTCT
DB	263	CCAACTCTT	TAGATACAT	TTCAGAGCGG	ATGTAACAGG	ATATTTAGGTAC
QY	683	GCAAGTCG	AGATTTTCA	CTCAATCCGT	710	
DB	203	ATGACGTG	GAGTTTACT	TACCCACCAG	T 176	

Search completed: August 27, 2003, 09:30:48
Job time : 3478.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:33 ; Search time 127.739 Seconds
(without alignments)
9607.814 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30
Sequence: 1 tccggatccagcgcctctgtttgatggct 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	30	100.0	30	6	AR157942	AR157942 Sequence
2	30	100.0	30	6	AX464425	AX464425 Sequence
3	28	93.3	1133	6	AR157944	AR157944 Sequence
4	28	93.3	1133	6	AX464427	AX464427 Sequence
5	21.6	72.0	119859	2	AP005865	AP005865 Oryza sat
6	21.6	72.0	125505	2	AP004177	AP004177 Oryza sat
7	21.6	72.0	316050	1	AX321859	AX321859 Nitrosomo
8	21.2	70.7	146405	2	AC143437	AC143437 Macaca mu
9	21.2	70.7	221887	2	AL158052	AL158052 Human DNA
10	21	70.0	1035	1	WSANSAGEN	X83689 W.succinoge
11	21	70.0	2505	1	WSDCUANSA	X89215 W.succinoge
12	21	70.0	3606	1	WSAJ2933	AX002933 Wolinella
13	20.8	69.3	4927	1	AE010450	AE010450 Methanopy
14	20.6	68.7	908	9	HSA340724	AJ340724 Homo sapi
15	20.6	68.7	13438	1	AE014767	AE014767 Bifidobac
16	20.6	68.7	172378	2	AC134210	AC134210 Rattus no
17	20.6	68.7	249207	2	AC102734	AC102734 Mus muscu
18	20.6	68.7	299950	1	AP005372	AP005372 Thermosyn
19	20.6	68.7	349980	6	AX492786	AX492786 Sequence
20	20.6	68.7	349980	6	AX553953	AX553953 Sequence
21	20.4	68.0	781	1	ECECODXXI	X98143 E.coli DNA
22	20.4	68.0	781	1	ECECOPRRI	X75452 E.coli (CTR
23	20.4	68.0	2200	1	ECDNAHSDM	M95666 Bacterioph
24	20.4	68.0	4343	7	PP1PHDDOC	AF262044 Arabidops
25	20.4	68.0	16228	8	F6113	AB025611 Arabidops
26	20.4	68.0	50015	8	AB025611	AL035682 Human DNA
27	20.4	68.0	89163	9	HS1009H6	AL298940 Mus muscu
28	20.4	68.0	137998	2	AL928940	AL929142 Mouse DNA
29	20.4	68.0	207823	10	AL929142	AL021937 Human DNA
30	20.2	67.3	173354	9	HS149A16	AC069248 Homo sapi
31	20	66.7	157493	9	AC069248	AC024157 Homo sapi
32	20	66.7	160667	9	AC024157	AC119169 Mus muscu
33	20	66.7	180413	2	AC119169	AC120733 Rattus no
34	20	66.7	254721	2	AC120733	AC133099 Mus muscu
35	20	66.7	255406	2	AC133099	AC126068 Rattus no
36	20	66.7	268801	2	AC126068	AC134916 Mus muscu
37	20	66.7	305735	2	AC134916	AC120175 Mus muscu
38	20	66.7	343010	2	AC120175	I64799 Sequence 5
39	19.8	66.0	6642	6	I64799	AX377877 Sequence
40	19.8	66.0	6840	6	AX377877	Y08110 H.sapiens m
41	19.8	66.0	6840	9	HS1R11	E13292 cDNA encodi
42	19.8	66.0	6843	6	E13292	I64800 Sequence 6
43	19.8	66.0	6843	6	E16800	AX098219 Sequence
44	19.8	66.0	6902	6	AX098219	AX336385 Sequence
45	19.8	66.0	6981	6	AX336385	

ALIGNMENTS

RESULT 1
AR157942
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AR157942
Sequence 1 from patent US 6251388.
AR157942
AR157942.1 GI:16219884

linear PAT 17-OCT-2001

Unknown.
Unclassified.
1 (bases 1 to 30)
Durdan,D.L.
Utilization of Wolinella succinogenes asparaginase to treat
diseases associated with asparagine dependence
Patent: US 6251388-A 1 26-JUN-2001;

Matches	28;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..119959
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSUNBB0037J12"
34012 a 26175 c 25976 g 33746 t 50 others

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 85.7%; Pred. No. 57;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCTCTGTTTGATGGCT 30
Db 79713 CGCCTCCTCTCTCTGTTTGATGGCT 79686

RESULT 7
EX321859
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EX321859
Nitrosomonas europaea ATCC 19718, complete genome; segment 4/10.
BX321859.1 GI:30138466
complete genome.
Nitrosomonas europaea ATCC 19718
Nitrosomonas europaea ATCC 19718
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1
Chain, P., Lamerdin, J., Larimer, F., Regala, W., Land, M., Hauser, L.,
Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D.,
Hommes, N., Whittaker, M. and Arp, D.
Complete Genome Sequence of the Ammonia-Oxidizing Bacterium and
Obligate Chemolithoautotroph Nitrosomonas europaea
J. Bacteriol. 185 (9), 2759-2773 (2003)
12700255
2 (bases 1 to 316050)
Larimer, F.
Direct Submission
Submitted (12-NOV-2002) Submitted on behalf of the Nitrosomonas
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov
Nitrosomonas genome consortium

REMARK
FEATURES
source
1..316050
/organism="Nitrosomonas europaea ATCC 19718"
/mol_type="genomic DNA"
/strain="ATCC 19718"
/db_xref="taxon:228410"
complement(639..1223)
/locus_tag="NE0846"
complement(639..1223)
/locus_tag="NE0846"
/locus_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD84757.1"
/db_xref="GI:30138467"
/translation="MIRNNSFYRIDPLKSPIKANPLRNPGESLNHLQNFILDGIV
PYFTANCFVLIAAEWRVWTSNPVIAICIAITLKIYKGRKEVKRIKL
GLAGLAVQFLERLRAGGSHIFHDIPKGFNLHVHSSGIIVETKLSKPDGRG
SKLYVDGNHILKNSALDRNFIITQVANSRWLRE"
order(complement(1038..1094),complement(951..1007))
/locus_tag="NE0846"
/note="2, probable transmembrane helices predicted by
TMHMM2.0"
complement(1351..2301)
/locus_tag="NE0847"
complement(1351..2301)
/locus_tag="NE0847"
/function="COGs COG0726"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD84758.1"
/db_xref="GI:30138468"
/translation="WKILPSLPSPAGKNSKLSIVYHRLPEPDLTGTGGGIAQPEKG

misc_feature
gene
CDS

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..119959
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSUNBB0037J12"
34012 a 26175 c 25976 g 33746 t 50 others

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 85.7%; Pred. No. 57;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCTCTGTTTGATGGCT 30
Db 109538 CGCCTCCTCTCTCTGTTTGATGGCT 109565

RESULT 6
AP004177/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AP004177
Oryza sativa (japonica cultivar-group) chromosome 2 clone
OJ1116 E03, *** SEQUENCING IN PROGRESS ***.
AP004177
GI:15718434
HTG; HTGS PHASE2.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OJ1116 E03
Published Only in Database (2001)
2 (bases 1 to 125505)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (20-SEP-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: this is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..125505
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="Genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clones="OJ1116 E03"
33795 a 27965 c 29679 g 34012 t 54 others

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 72.0%; Score 21.6; DB 2; Length 125505;

Thu Aug 28 09:00:58 2003

```

RESULT 8
AC143437
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC143437 146405 bp DNA linear HTG 09-APR-2003
Macaca mulatta clone CH250-267L24, *** SEQUENCING IN PROGRESS ***.
AC143437
HTG; HTGS PHASE2; HTGS_PGI.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 146405)
Ceuroso,M. and Milosavljevic,A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo,R. and Gusfield,D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 146405)
Milosavljevic,A., Sodergren,E., Ceuroso,M., Li,B., Jackson,A.R.,
Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L.,
Amaratunge,H.C., Are,J.R., Avelle,M., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Egan,A., Earhart,T., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsif,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondei,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogih,M., Okuwonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleznyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
3 (bases 1 to 146405)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
-----

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LBPH
Center clone name: CH250-267L24
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; inf% of reads
Consensus quality: 4889 bases at least Q40
Consensus quality: 5926 bases at least Q30
Consensus quality: 6870 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: The contigs are based on the application draft_data.html)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 146405: contig of 146405 bp in length.
* Location/Qualifiers
* 1..146405
* /organism="Macaca mulatta"
* /mol_type="genomic DNA"
* /db_xref="taxon:9544"
* /clone="CH250-267L24"
* 1..146405
* /note="assembly_name:CH250-267L24.1B
* CONFIDENCE=0.67"
BASE COUNT 2409 a 2871 c 2652 g 2518 t 135955 others
ORIGIN
Query Match 70.7%; Score 21.2; DB 2; Length 146405;
Best Local Similarity 88.5%; Pred. No. 88;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCCGGATCCAGCGCTCTGTTTGTAT 26
Db 95612 TCCGGAACCTGGGCTCTGTTTGTAT 95637
-----
RESULT 9
AL158052
LOCUS
DEFINITION
AL158052 Human DNA sequence from clone RP11-570J20 on chromosome
9q33.1-34.12, complete sequence.
ACCESSION
VERSION AL158052.10 GI:12964258
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 221887)
Direct Submission
Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 19, 2001 this sequence version replaced gi:12657095.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
-----
TITLE
JOURNAL
AUTHORS
COMMENT

```

Submitted (05-JAN-1995) K.H. Roehm, Physiologische Chemie der Philipps-Univ., Karl-von-Frisch-Str.1, D-35033 Marburg (Lahn), FRG

```

/allele= ucua
/codon start=1

```

```

/organism="Wolinella succinogenes"
/mol_type="genomic DNA"
/db_xref="taxon:844"
/clone_lib="lambda EMBL3"
1..1068
/gene="dcuA"
<1..1068
/gene="dcuA"
/codon_start=1

```



```

/transl_table=11
/product="C4-dicarboxylate membrane transporter"
/protein_id="CAA61502.1"
/db_xref="GI:895918"
/db_xref="SWISS-PROT:O34245"
/translation="KLESQPKHITFLAPVVTYFMTLLAGTGFATFSLPVIKVAKE
QGVKSVLEVCGRGLSIAITASPVSAVFMGRVDELGVGLQLLAICTPTTFCG
MITATICNLFNDLQDPYQERVAKGMVKRGQTQYVAKGAKLSVILFLLGILAVV
FYATISKVNLQNPQIVGRDSAINMFMLTATLIAMFPAKIDTDAVLKASTKSGMTA
CTICVLGAWLQDTFVSNINDIKALSGGILEGHPWMLAIILFFASMLLYSQATAKAL
IPAAVPLSDVPVSLIASPAASGLFVLPTPTLLAAVQMDTGTSTRICKFVFNHPFV
PGVLAIGISVALGFIIVAPILL"
misc_feature
1069..1188
/note="intergenic region"
terminator
1079..1102
/note="potential rho-independent terminator for dcua"
RBS
1171..1177
/note="putative"
gene
1189..2181
/genes="ansa"
1189..2181
/genes="ansa"
CDS
/EC_number="3.5.1.1"
/codon_start=1
/transl_table=11
/product="asparaginase"
/protein_id="CAA61503.1"
/db_xref="GI:895919"
/db_xref="SWISS-PROT:P50286"
/translation="WAKPOVITLGTGGTAGSGSSVSKSSAGAVTVDKLLAAVPAI
NLTATKQSIQSGQWGTGKVLKRVNLLAOKETRAVITHTGDTMEETAFP
LNTVKSQRPVLVGNMRGSSMSADGPNLYNVAINKASTNKGVVIVMDEIHA
AREATKLTAVNAFASPNTGKIGYVYKGVYFTQSVRPHTLASEPDISKIELPRV
DILYARQDTEVDVNAALQAGAKIIGHAGNGNPNPPLTONAKAAGSGVYVAKRSR
VSGSGTTPQAEYDDKKLGFVATESLNPQARVLLMLALTKTSDREAIQKIFSTY"
2186..2219
terminator
/note="potential rho-independent terminator for ansa"
BASE COUNT 594 a 690 c 553 g 668 t
ORIGIN
70.0%; Score 21; DB 1; Length 2505;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCCTCTGTTTGTGGCT 30
|||||
Db 1174 AGCGCCTCTGTTTGTGGCT 1194

RESULT 12
WSAJ2933 3606 bp DNA linear BCT 05-OCT-2000
LOCUS Wolinella succinogenes aspa, dcua genes and partial ansa gene.
DEFINITION
AJ002933
VERSION AJ002933.1 GI:2644958
KEYWORDS ansa gene; aspa gene; aspartate ammonia-lyase; C4-dicarboxylate membrane transporter; dcua gene; L-asparaginase.
SOURCE Wolinella succinogenes
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Wolinella.

REFERENCE
1 Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.
Transport of C(4)-dicarboxylates in Wolinella succinogenes
J. Bacteriol. 182 (20), 5757-5764 (2000)
20461222
MEDLINE
11004174
PUBMED
REFERENCE
2 (bases 1 to 3606)
Ullmann, R.
Direct Submission
AUTHORS Ullmann, R.
TITLE Transport of C(4)-dicarboxylates in Wolinella succinogenes
JOURNAL J. Bacteriol. 182 (20), 5757-5764 (2000)
MEDLINE 20461222
PUBMED 11004174
REFERENCE
2 (bases 1 to 3606)
Ullmann, R.
Direct Submission
AUTHORS Ullmann, R.
TITLE Transport of C(4)-dicarboxylates in Wolinella succinogenes
JOURNAL J. Bacteriol. 182 (20), 5757-5764 (2000)
MEDLINE 20461222
PUBMED 11004174

Query Match 70.0%; Score 21; DB 1; Length 3606;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCCTCTGTTTGTGGCT 30
|||||
Db 3235 AGCGCCTCTGTTTGTGGCT 3255

RESULT 13
AE010450
LOCUS Methanopyrus kandleri AV19 section 149 of 157 of the complete genome..
DEFINITION
AE010450
ACCESSION AE010450 AE009439

COMMENT
Related sequence: X89215.
Location/Qualifiers
1..3606
/organism="Wolinella succinogenes"
/mol_type="genomic DNA"
/db_xref="taxon:844"
400..1815
/genes="aspa"
400..1815
/genes="aspa"
CDS
/codon_start=1
/transl_table=11
/product="aspartate ammonia-lyase"
/protein_id="CAA05764.1"
/db_xref="GI:2644959"
/db_xref="SPTREMBL:O34244"
/translation="MASTRIEHDLIGDLEVSNDYCYGQVOTARAKENFHIISGVLLSSMP
TFIASLAKVKKAAALANFELGLLDEYKIKNAICEACDLIITAGKHDFVVDALQGGAGT
STNNANEVIANVALELMGHKKGEYKHPNNHNSQSTNDAYPTAFRVALYKELIE
LTVSMGILRDSFAKAEFKDIIKMGRTQLQDAVPMTLGQEFKYAVMMTDTDRVLE
ARNLREMNLGATAIGTINSHPDYATLVKKLQEVTVGRPFITASDLVEATQDTGAYV
QISGLKRVSTKVSICNDLRLLSSGRAGLSINLPMPQSGSIIMPQKVPVPIPVV
NOQYHVIGADMVTIIASGGQLQANVFPFVIAFSLFSSINMMRAFEFLAEKCVNGI
TANADNCRDYLKISIGLITNALNPYIGTSTNSVAKAELSGSKSVHDIIVLERKLLSKEE
LEEIIRPNMIKPKVVTLTCK"
1828..3129
/genes="dcua"
1828..3129
/genes="dcua"
CDS
/codon_start=1
/transl_table=11
/product="C4-dicarboxylate membrane transporter"
/protein_id="CAA05765.1"
/db_xref="GI:2644960"
/db_xref="SWISS-PROT:O34245"
/translation="MLVFELVVLAAIFLQVRLLGGTAIGYAGGLGMVLCGLGLKPG
SIFDVIILMSVIAAAMQVAGGLDYLHLAEKLLRSQPKHITFLAPVVTFTMILL
AGTGHYATFSLPVIKVAKEQGVKPSVPLSIAVASQIAITASPVSAVFMGRVLE
LGVYQLLAICTPTTFCGMITATLIAMFPAKIDTDAVLKASTNKGVVIVMDEIHA
AKPKALSVILFVLGILAVVYATVATSKVGLIQNPVGRDSAINMFMLTATLIAMF
AKIDTDAVLNASTFKSGMTACICVLGAWLQDTFVSNINDIKGFAAGILEGHPWMLA
ITLFFASMLLYSQATAKALIPAAVPLSDVPVSLIASPAASGLFVLPTPTLLAAVQ
MDTGTSTRICKFVFNHPFVPGVLAIGISVALGFIIVAPILL"
3250..3606
/genes="ansa"
3250..3606
/genes="ansa"
CDS
/codon_start=1
/transl_table=11
/product="L-asparaginase"
/protein_id="CAA05766.1"
/db_xref="GI:2644961"
/translation="WAKPOVITLGTGGTAGSGSSVSKSSAGAVTVDKLLAAVPAI
NLTATKQSIQSGQWGTGKVLKRVNLLAOKETRAVITHTGDTMEETAFP"
NLNLTGKQKPVVLVGM"
BASE COUNT 919 a 973 c 806 g 908 t
ORIGIN
70.0%; Score 21; DB 1; Length 3606;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCCTCTGTTTGTGGCT 30
|||||
Db 3235 AGCGCCTCTGTTTGTGGCT 3255

RESULT 13
AE010450
LOCUS Methanopyrus kandleri AV19 section 149 of 157 of the complete genome..
DEFINITION
AE010450
ACCESSION AE010450 AE009439

```


/translation="MTERIDATNATEHMADAESTAGSKAAVRGVEETREYDVFDTVD
DPFGPCLAEALIMVAQDQTELAIALDNTVVEEALQALHDDYDEGSEGRVPR
RQFLRLTARGHQEASRAAFQFVSFYTDQTLRQALEALAVAYRQPVTRAQV
AALRGVNSGVIRSLTVRGLIKEDGVDPESRAALTVSEFFLQMGUKSLDELPSLAP
FUPSAEMTQSKSGDITVIS"
5836..6663

gene
CDS

/gene="BL1373"
/note="COG family: ADP-ribose pyrophosphatase; PFAM ID:

```
NUDIX"
/codon_start=1
/transl_table=11
/product="narrowly conserved hypothetical protein possibly
in the MutT/Nudix family"
/protein_id="RAN25174.1"
/db_ref="RAN25174.1"
/db_ref="RAN25174.1"
```

/w/.../...-u/...:43326643"
 Translation: "MGFNGVNSRSPAPQGVSVVILALPGKPCGNSAADSNGAANS
 AGHGRWRLSLVKVRQVFLGWCALPGDLSDRLSLEQSAVALESTTDLHPVYLEQLYT
 FOSFTRPSHGGLPMVSVYVWALVGTGAAGFDGNWVPFDELPELAHQRIIDYA
 LLRLRSKTEYSDVATRLGFTTQLRQHUYVEAIGETDILANFRMLASGLDGTG
 GREGROGRQATPVYRYTPKRPVENGTPHESGDGDIVELSRQGERDALDALIPSAI"
 5724 - 8004

CDS

/gene="nadaA"
/note="NadaA; COG family: quinolate synthase; PFAM_ID:
nadaA"

```
codon_start=1
/transl_table=11
product="guinolate synthetase"
protein_id="AA25175.1"
db_xref="GI:23326646"
/
```

translating="M7ALSTAP5VSEI1AKLGAQSCD7AGL7QD7PHW7D7V7GPG7
K7ASMD7R7P7ANAP7Q7V7L7STAP7S7R7K7S7D7E7E7I7Q7R7I7TAK7Q7R7K7S7K7L7I7GH7F7Q7R7D7E7I7
7V7G7D7S7Q7LAK7N7A7T7E7D7V7FC7G7V7H7F7AE7AT7D7L7S7T7E7Q7S7V7L7N7L7S7AG7S7MA7
7M7AN7D7I7Q7E7C7O7O7L7G7E7C7I7C7D7P7D7S7Q7L7I7P7V7T7V7N7S7A7A7K7A7C7R7G7N7I7V7C7T7S7
7TAK7AL7H7S7W7A7R7K7G7V7L7F7F7D7H7L7G7S7TAR7G7A7M7I7P7L7S7E7M7L7P7D7F7Q7A7G7A7D7P7A7V7
7K7A7M7L7K7M7G7S7C7H7R7V7T7E7Q7E7R7A7K7V7G7Y7K7V7I7P7E7C7S7M7V7D7A7D7A7G7T7S7TAY7I7
7E7I7AN7A7P7A7S7A7I7V7G7E7I7N7L7N7L7A7Q7I7P7D7C7L7D7P7C7S7T7M7Y7I7H7P7AY7L7AW7AL7
7N7E7A7G7N7V7N7K7I7V7T7E7D7E7A7K7V7A7L7E7M7L7R7H7P7"

SDS

```

0024: 5743
gene="nadB"
note="guinolate synthetase B; NadB; COG family:
spartate oxidase; PFAM_ID: pyr_redox; PFAM_ID:
AD_binding_2; PFAM_ID: succ_DH_flav_C"
codon_start=1
transl_table=11

```

```
product="L-aspartate oxidase"  
protein_id="AAN25176.1"  
db_xref="GI:23326647"
```

21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 54

ene

27.10620
name="nadC"
note="quinolate phosphoribosyltransferase
decarboxylating"; Qaprtase; NadC; COG family:
Icotinate-nucleotide pyrophosphorylase; PFAM ID:

```

QRPase_N; PFAM_ID: QRPase"
/codon_start=1
/transl_table=11
product="probable nicotinate-nucleotide pyrophosphorylase
[carboxylating]"
protein_id="AAN25177.1"
db_xref="GI:2326648"
translation="MLTSHIRVAVEAALADAPYGDITCETTPADETSAGHITARE
GNGVMSGIDVPAFAAFQANPVTVTAIKOGERQFGQILATVEGVDLLTLTAERIALN
ADVDVAVYADDYDGDPVTPRRYRTRI VDTKRTPLGLRPFKY
TQRMSSGATWTAAFDVAVNGLDGLAALAAARGDLALAI RHVREQVGHGTHIEVEVDR
QVCGGHNHRRYGLSDAVKMNKHLAALAAARGDLALAI RHVREQVGHGTHIEVEVDR
DQIPFATAGGADDTIMLNDPFLSDTTRRGVELIDGKAIVASGNSMLRVPAAVATGVD
IISVGALTHSVRSIDLGLDWN"
0.0623..11870
'gene="BL1377"
0.0623..11870
'gene="BL1377"
note="COG family: cysteine sulfinatase desulfinate/cysteine
desulfurase and related enzymes; PFAM_ID: aminotran_5"
/codon_start=1
/transl_table=11

```

Query Match

Best Local Similarity 85.2%; Pred. No. 1.7e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 3 CGGATCCAGCGCCTCTGTTTGGATGCC 29
 |||||
 1327 CGGATCCAGCGACTCTGCTCGATGCC 1353

Arch completed: August 27, 2003, 02:34:23
 Time : 131.739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:23 ; Search time 11.6853 Seconds
(without alignments)
6930.345 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tcggatccagcgcctctgtttgatggct 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	20	AA03472 Wolinella succinog
2	30	100.0	30	21	W. succinogenes L-
3	28	93.3	1133	20	AA03474 Wolinella succinog
4	28	93.3	1133	21	W. succinogenes as
5	20.6	68.7	349980	24	AB081848 Bifidobacterium lo
6	19.8	66.0	6840	24	ABX84742 Human cDNA differe
7	19.8	66.0	6840	24	AAU88419 Pain regulated cDN
8	19.8	66.0	6843	18	AAU84509 Human LDL receptor

ALIGNMENTS

RESULT 1

AA03472

ID AA03472 standard; DNA; 30 BP.

XX AA03472;

AC AA03472;

XX 30-APR-1999 (first entry)

XX Wolinella succinogenes L-asparaginase PCR primer 1.

XX L-asparaginase; amplification; treatment; disease; asparagine depletion;

XX malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;

XX autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;

XX covalent modification; acylation; pharmacokinetic; immunogenic; spleen;

XX hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;

XX lymphocyte marker distribution; hepatotoxicity; PCR primer; ss.

XX Synthetic.

OS Wolinella succinogenes.

XX WO9856410-A1.

PD 17-DEC-1998.

XX 09-JUN-1998; 98WO-US11905.

XX 09-JUN-1997; 97US-0049085.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

Human late stage o
Human bladder can
cDNA encoding huma
Human benign prost
Kidney cancer rela
Human breast cance
Novel human cDNA s
Lung cancer relate
Partial genomic DN
Novel human polynu
cDNA encoding nove
DNA encoding novel
E. coli DNA for ce
Drosophila melanog
Genomic sequence #
DNA encoding novel
Drosophila melanog
Drosophila melanog
Novel human coding
Drosophila melanog
Drosophila melanog
Aspergillus fumiga
Aspergillus fumiga
Rabbit LDL recepto
Arabidopsis thalia
Corn ear-derived p
Human secreted pro
Human musculoskele
cDNA encoding nove
Arabidopsis thalia
Arabidopsis thalia
Human cDNA encodin
Human novel polynu
Human polynucleoti
DNA encoding human

9 19.8 66.0 6902 22 AAF98724
10 19.8 66.0 6906 25 AAC50992
11 19.8 66.0 6981 24 ABS76520
12 19.8 66.0 6981 24 ABK64741
13 19.8 66.0 6981 24 ABL68557
14 19.8 66.0 7450 25 ABT31940
15 19.4 64.7 73308 24 ABL6966
16 19.4 64.7 4766 19 AAV42563
17 18.6 62.0 352 22 AAF65999
18 18.4 61.3 574 23 ABX41664
19 18.4 61.3 1317 23 AAS88555
20 18.4 61.3 1365 23 AAS2663
21 18.4 61.3 1673 23 ABL18335
22 18.4 61.3 2407 23 ABK42208
23 18.4 61.3 2688 23 AAS79968
24 18.4 61.3 4542 23 ABL09869
25 18.4 61.3 5347 23 ABH11058
26 18.4 61.3 6630 24 ABN59745
27 18.4 61.3 6835 23 ABL09868
28 18.4 61.3 7875 23 ABL18334
29 18.4 61.3 12816 23 ABL20852
30 18.2 60.7 2621 25 ABT17801
31 18.2 60.7 2621 25 ABT19615
32 18.2 60.7 6961 18 AAT84508
33 18.2 60.7 162 21 AAC50042
34 18 60.0 277 25 ABX84590
35 18 60.0 455 21 AAC31008
36 18 60.0 1333 22 ABL37539
37 18 60.0 1333 25 ABX60327
38 18 60.0 1686 21 AAC42929
39 18 60.0 1686 24 ABL214235
40 18 60.0 1695 22 AAS26188
41 18 60.0 1695 25 ABX73529
42 18 60.0 2380 24 ABZ11903
43 18 60.0 2452 24 ABZ11904
44 18 60.0 5210 24 ABK52202
45 18 60.0 5210 24 ABK52202

XX WPI; 1999-070295/06.
 XX Treating diseases responsive to asparagine depletion - especially
 PT cancer and auto-immune diseases, using a Wolinella succinogenes
 PT asparaginase
 XX
 PS Example 4; Fig 1; 58pp; English.
 XX
 CC This sequence is a PCR primer used in the isolation and amplification of
 CC an L-asparaginase gene from Wolinella succinogenes which is used in a
 CC method for the treatment of diseases that respond to asparagine depletion
 CC This asparaginase and its analogues, are particularly used to treat (i)
 CC malignant diseases, particularly where haematological and specifically
 CC lymphoma, leukaemia and myeloma (in either chronic or acute phases) and
 CC (ii) autoimmune diseases, specifically rheumatoid arthritis, systemic
 CC lupus erythematosus and acquired immune deficiency syndrome, in humans
 CC or other mammals. The asparaginase can also be used for the covalent
 CC modification of biologically active proteins by acylation which is used
 CC to alter pharmacokinetic and immunogenic properties of biologically
 CC active proteins, and can be applied to any protein used to treat animal,
 CC human or plant diseases. The asparaginase can be used to treat patients
 CC who have become hypersensitive to other microbial asparaginases. The
 CC asparaginase has reduced toxicity and reduced tendency to cause
 CC immunosuppression or allergy, particularly it has no effect on spleen and
 CC thymus histology or lymphocyte marker distribution and no hepatotoxicity.
 XX
 SQ Sequence 30 BP; 3 A; 9 C; 8 G; 10 T; 0 other;
 Query Match 100.0%; Score 30; DB 20; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.0006;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30
 Db 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30
 RESULT 2
 AAC62510
 ID AAC62510 standard; DNA; 30 BP.
 XX AAC62510;
 AC AAC62510;
 XX
 DT 07-FEB-2001 (first entry)
 XX
 DE W. succinogenes L-asparaginase PCR primer #1.
 XX
 KW Asparaginase; glutaminase; autoimmune disease; graft versus host disease;
 KW leukaemia; PCR primer; ss.
 XX
 OS Wolinella succinogenes.
 XX
 PN WO200059533-A1.
 XX
 PD 12-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US07981.
 XX
 PR 02-APR-1999; 99US-0127662.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.
 XX
 PI Durden DL;
 XX
 DR WPI; 2000-664963/64.
 XX
 CC Treating or preventing graft versus host and autoimmune diseases, e.g.
 PT diabetes, rheumatoid arthritis, that responds to asparagine or
 PT glutamine depletion comprises administering an asparaginase or
 PT glutaminase -
 XX
 XX Example 4; Fig 1; 72pp; English.

XX The present invention relates to a novel method of using the recombinant
 CC microbial enzymes asparaginase and glutaminase to treat autoimmune
 CC diseases and leukaemia. The diseases likely to respond to such treatment
 CC include graft versus host disease, rheumatoid arthritis, systemic lupus
 CC erythematosus, autoimmunity, collagen vascular diseases, AIDS,
 CC osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple
 CC sclerotic diseases, primary biliary cirrhosis, rheumatic fever, inflammatory
 CC bowel diseases; primary biliary cirrhosis, chronic active hepatitis,
 CC glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves,
 CC disease. The present sequence is a PCR primer used in the amplification
 CC of the W. succinogenes asparaginase sequence.
 XX
 SQ Sequence 30 BP; 3 A; 9 C; 8 G; 10 T; 0 other;
 Query Match 100.0%; Score 30; DB 21; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.0006;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30
 Db 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30
 RESULT 3
 AAX03474
 ID AAX03474 standard; DNA; 1133 BP.
 XX AAX03474;
 AC AAX03474;
 XX
 DT 30-APR-1999 (first entry)
 XX
 DE Wolinella succinogenes L-asparaginase DNA.
 XX
 KW L-asparaginase; amplification; treatment; disease; asparagine depletion;
 KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW covalent modification; acylation; pharmacokinetic; immunogenic; spleen;
 KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;
 KW lymphocyte marker distribution; hepatotoxicity; ss.
 XX
 OS Wolinella succinogenes.
 XX
 PN WO9856410-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 09-JUN-1998; 98WO-US11905.
 XX
 PR 09-JUN-1997; 97US-0049085.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.
 XX
 PI Durden DL;
 XX
 DR WPI; 1999-070295/06.
 XX
 PT Treating diseases responsive to asparagine depletion - especially
 PT cancer and auto-immune diseases, using a Wolinella succinogenes
 PT asparaginase
 XX
 PS Example 5; Fig 6; 58pp; English.
 XX
 CC This sequence is a novel L-asparaginase gene from Wolinella succinogenes
 CC which is used in a method for the treatment of diseases that respond to
 CC asparagine depletion. This asparaginase and its analogues, are
 CC particularly used to treat (i) malignant diseases, particularly where
 CC haematological and specifically lymphoma, leukaemia and myeloma (in
 CC either chronic or acute phases) and (ii) autoimmune diseases,
 CC specifically rheumatoid arthritis, systemic lupus erythematosus and
 CC acquired immune deficiency syndrome, in humans or other mammals. The
 CC asparaginase can also be used for the covalent modification of
 CC biologically active proteins by acylation which is used to alter

CC pharmacokinetic and immunogenic properties of biologically active
 CC proteins, and can be applied to any protein used to treat animal, human
 CC or plant diseases. The asparaginase can be used to treat patients who
 CC have become hypersensitive to other microbial asparaginases. The
 CC asparaginase has reduced toxicity and reduced tendency to cause
 CC immunosuppression or allergy, particularly it has no effect on spleen and
 CC thymus histology or lymphocyte marker distribution and no hepatotoxicity.

XX Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 other;

Query Match 93.3%; Score 28; DB 20; Length 1133;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CGGATCCAGCGCTCTGTTTGTATGGCT 30
 DB 96 CGGATCCAGCGCTCTGTTTGTATGGCT 123

RESULT 4
 AAC62512
 ID AAC62512 standard; DNA; 1133 BP.

XX AAC62512;

XX 07-FEB-2001 (first entry)

DE W. succinogenes asparaginase-specific DNA insert.

XX Asparaginase; glutaminase; autoimmune disease; graft versus host disease;
 KW leukaemia; DNA insert; ds.

XX Wolinella succinogenes.

XX WO200059533-A1.

XX 12-OCT-2000.

XX 24-MAR-2000; 2000WO-US07981.

XX 02-APR-1999; 99US-0127662.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DI;

XX WPI; 2000-664963/64.

XX Treating or preventing graft versus host and autoimmune diseases, e.g.

PT diabetes, rheumatoid arthritis, that responds to asparagine or

PT glutamine depletion comprises administering an asparaginase or

PT Glutaminase -

XX Example 5; Fig 6; 72pp; English.

XX The present invention relates to a novel method of using the recombinant
 CC microbial enzymes asparaginase and glutaminase to treat autoimmune
 CC diseases and leukaemia. The diseases likely to respond to such treatment
 CC include graft versus host disease, rheumatoid arthritis, systemic lupus
 CC erythematosus, autoimmunity, collagen vascular diseases, AIDS,
 CC osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple
 CC sclerosis, sclerosing pancreatitis, rheumatic fever, inflammatory
 CC bowel diseases, primary biliary cirrhosis, chronic active hepatitis,
 CC glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves'
 CC disease. The present invention is DNA insert used to demonstrate the
 CC method of the invention.

XX Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 other;

Query Match 93.3%; Score 28; DB 21; Length 1133;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CGGATCCAGCGCTCTGTTTGTATGGCT 30
 DB 96 CGGATCCAGCGCTCTGTTTGTATGGCT 123

RESULT 5
 ABQ81848

ID ABQ81848 standard; DNA; 349980 BP.

XX ABQ81848;

XX 19-NOV-2002 (first entry)

DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1104.

XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KW rotavirus; food composition; pharmaceutical composition; gene; ds.

XX Bifidobacterium longum.

OS Synthetic.

XX EPI227152-A1.

XX 31-JUL-2002.

XX 30-JAN-2001; 2001EP-0102050.

XX 30-JAN-2001; 2001EP-0102050.

XX (NEST) SOC PROD NESTLE SA.

XX WPI; 2002-668397/72.

XX Novel polynucleotide comprising Bifidobacterium genome sequence useful
 PT as a probe or primer for detecting and/or identifying Bifidobacterium
 PT longum in a biological sample -

XX Disclosure; SEQ ID 1104; 80pp; English.

XX The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding
 CC a fusion protein, comprising a sequence selected from 1097 sequences
 CC given in ABP65258 to ABP66354 ligated in frame to a polynucleotide
 CC encoding a heterologous polypeptide. (I) has antidiarrheic and
 CC antibacterial activities, and can be used as an inhibitor of Salmonella.
 CC (I) (which is a probe) is useful for the detection and/or identification
 CC of Bifidobacterium longum in a biological sample. A carrier containing
 CC the lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM I-2618)
 CC can be used for preventing and/or treating diarrhoea brought about by
 CC pathogenic bacteria and/or rotavirus. The carrier is a food composition
 CC selected from milk, yogurt, curd, cheese, fermented milks, milk based
 CC fermented products, ice-creams, fermented cereal based products, milk
 CC based powders, infant formula, pet food or a pharmaceutical composition
 CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (I) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the Bifidobacterium genome. ABQ81844 to ABQ81850 represent
 CC Bifidobacterium related nucleotide sequences given in the Sequence
 CC Listing from the present invention but not mentioned further within the
 CC specification.

XX N.B. The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied by the
 CC European Patent Office.

XX Sequence 349980 BP; 69195 A; 106952 C; 106128 G; 67705 T; 0 other;

Query Match 68.7%; Score 20.6; DB 24; Length 349980;
 Best Local Similarity 85.2%; Pred. No. 41;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CGGATCCAGCGCCTCTGTTTGATGGC 29
 |||||
 Db 202894 CGGATCCAGCGCCTCTGTTTGATGGC 202920

RESULT 6
 ID ABL88419 standard; cDNA; 6840 BP.
 AC ABL88419;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #1313.
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 03-OCT-2001; 2001WO-US30821.
 XX
 PF 03-OCT-2000; 2000US-237189P.
 XX
 PR (GENE-) GENE LOGIC INC.
 PA
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX
 PS Claim 1; SEQ ID No 1313; 114pp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GCA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GCA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile

CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection, and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 6840 BP; 1679 A; 1744 C; 1831 G; 1586 T; 0 other;

Query Match 66.0%; Score 19.8; DB 24; Length 6840;
 Best Local Similarity 91.3%; Pred. No. 53;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2514 ATCCAGCGCCTCTGTTTGATGG 2536

RESULT 7

ABL88419

ID ABL88419 standard; cDNA; 6840 BP.

XX

AC ABL88419;

XX

DT 16-MAY-2002 (first entry)

XX

DE Pain regulated cDNA sequence 62.

XX

KW Pain; analgesic; gene therapy; neurological disorder;

XX

KW neurodegenerative disease; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200212338-A2.

XX

PD 14-FEB-2002.

XX

PF 03-AUG-2001; 2001WO-EP09011.

XX

PR 03-AUG-2000; 2000DE-1037759.

XX

XX (CHEF) GRUENENTHAL GMBH.

XX

PI Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;

XX

DR WPI; 2002-257469/30.

XX

DR P-PSDB; ABB85016.

XX

PT Identifying pain-regulating compounds, useful for treating chronic pain
 PT and for diagnosis, by measuring binding of compounds to specific
 PT peptides and proteins -

XX

PS Claim 1; Fig 36; 213pp; German.

XX

CC The invention relates to identifying pain-regulating substances (A)
 CC comprises (i) incubating a test substance with a cell (or preparation
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
 CC either binding of the test substance to (B) or some functional parameter
 CC that is altered by this binding. The method is useful for identifying
 CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B,
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polynucleotide of the invention.

XX SQ Sequence 6840 BP; 1679 A; 1744 C; 1831 G; 1586 T; 0 other;
 Query Match 66.0%; Score 19.8; DB 24; Length 6840;
 Best Local Similarity 91.3%; Pred. No. 53;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28
 |||||
 Db 2514 ATCCAGCGCCTCTGTTTGAATGG 2536

RESULT 8
 AAT84509
 ID AAT84509 standard; cDNA to mRNA; 6843 BP.
 XX
 AC AAT84509;
 XX
 DT 01-DEC-1997 (first entry)
 XX
 DE Human LDL receptor analogue cDNA.
 XX
 KW Low density lipoprotein receptor analogue; LDL receptor;
 KW lipoprotein; arteriosclerosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 81..6725
 FT Location/Qualifiers
 FT /*tag= a
 FT /note= "Claim 3"
 FT 81..164
 FT sig_peptide
 FT /*tag= b
 FT 165..6722
 FT mat_peptide
 FT /*tag= c
 FT
 XX EP773290-A2.
 XX
 PD 14-MAY-1997.
 XX
 PF 08-OCT-1996; 96BP-0116108.
 XX
 PR 24-APR-1996; 96JP-0102451.
 PR 09-OCT-1995; 95JP-0261440.
 XX
 PA (KOWA) KOWA CO LTD.
 XX
 PI Arai K, Iwasaki A, Saito Y, Yamazaki H;
 XX
 DR WPI; 1997-261317/24.
 DR P-PSDB; AAW26357.
 XX
 PT DNA encoding LDL receptor analogue proteins - for research into
 PT arteriosclerosis
 XX
 PS Claim 3; Page 97-103; 104pp; English.
 XX
 CC This full-length cDNA sequence, of which the coding sequence is
 CC claimed, codes for a novel human low density lipoprotein
 CC (LDL) receptor family member (AAW26357) which participates in
 CC lipoprotein metabolism and which is a critical factor that triggers
 CC the onset of arteriosclerosis. It was identified by screening a
 CC human brain cDNA library with a partial cDNA clone. A vector
 CC comprising the coding sequence can be used in a claimed method to
 CC produce LDL receptor analogue protein in transformant cells. A
 CC cDNA clone (see AAW84508) coding for a novel rabbit LDL receptor
 CC (AAW26356) is also claimed. It is expected that, through analysis of
 CC the protein, details of lipoprotein metabolism mediated by the
 CC membrane LDL receptor will be elucidated, and the pathology of
 CC abnormal lipid metabolism which triggers the onset and progression
 CC of arteriosclerosis will be clarified.
 XX
 SQ Sequence 6843 BP; 1682 A; 1744 C; 1831 G; 1586 T; 0 other;

Query Match 66.0%; Score 19.8; DB 18; Length 6843;
 Best Local Similarity 91.3%; Pred. No. 53;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28
 |||||
 Db 2514 ATCCAGCGCCTCTGTTTGAATGG 2536

RESULT 9
 AAF98724
 ID AAF98724 standard; DNA; 6902 BP.
 XX
 AC AAF98724;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE Human late stage ovarian tumour polynucleotide marker 39.
 XX
 KW Human; ovarian cancer; identification; detection; characterisation;
 KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200118542-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24199.
 XX
 PR 03-SEP-1999; 99US-0152547.
 PR 16-MAR-2000; 2000US-0190347.
 PR 21-MAR-2000; 2000US-0191321.
 PR 31-MAY-2000; 2000US-0208382.
 PR 20-JUL-2000; 2000US-0220467.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lee J, Thompsho P, Lillie J;
 XX
 DR WPI; 2001-211428/21.
 XX
 PT Detection, assessment, prevention and therapy of ovarian cancer,
 PT comprises detecting changes in the expression of a variety of markers -
 XX
 PS Claim 1; Page 1189-1191; 1199pp; English.
 XX
 CC The present invention describes a method for assessing whether a patient
 CC is afflicted with ovarian cancer by comparing: (1) the expression of a
 CC marker (I) (see AAF98594 to AAF98730), in a patient sample; and (2) the
 CC normal level of expression of (I) in a control non-ovarian cancer
 CC sample, where a significant difference between the level of expression
 CC in (a) and (b) is an indication that the patient is afflicted with
 CC ovarian cancer. (I) have cytostatic activities and can be used in
 CC antisense gene therapy. The method, compositions and kits from the
 CC present invention can be used for: (1) assessing and treating ovarian
 CC cancer; (2) making isolated hybridoma, which produces an antibody useful
 CC for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a
 CC patient. AAF98573 to AAF98593 represent human kinase marker primers and
 CC probes which are used in the exemplification of the present invention.
 XX
 SQ Sequence 6902 BP; 1719 A; 1747 C; 1837 G; 1599 T; 0 other;

Query Match 66.0%; Score 19.8; DB 22; Length 6902;
 Best Local Similarity 91.3%; Pred. No. 53;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28
 |||||
 Db 2510 ATCCAGCGCCTCTGTTTGAATGG 2532

```
RESULT 10
ACCS0992
ID ACCS0992 standard; cDNA; 6906 BP.
XX
AC ACCS0992;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:77.
XX
Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX
Homo sapiens.
XX
WO2003003906-A2.
XX
16-JAN-2003.
XX
03-JUL-2002; 2002WO-US21338.
XX
03-JUL-2001; 2001US-302814P.
XX
03-AUG-2001; 2001US-310099P.
XX
08-NOV-2001; 2001US-343705P.
XX
13-NOV-2001; 2001US-350666P.
XX
12-APR-2002; 2002US-372466P.
XX
(EOSB-) EOS BIOTECHNOLOGY INC.
XX
Mack DH, Aziz N;
XX
WPI; 2003-201532/19.
XX
P-PSDB; ABR48181.
XX
Detecting a bladder cancer-associated transcript in a cell from a
PT
patient, comprises contacting a biological sample from the patient with
PT
a bladder cancer-associated polynucleotide or antibody
PT
Claim 6; Page 258-259; 307pp; English.
XX
The present invention describes a method for detecting a bladder cancer-
CC
associated transcript in a cell from a patient. The method comprises
CC
contacting a biological sample from the patient with a polynucleotide
CC
that selectively hybridizes to a sequence that is 80 % identical to a
CC
table of sequences (see ACCS0951 to ACCS1059). ACCS0951 to ACCS1059
CC
encode the human bladder cancer-associated proteins given in ABR48146 to
CC
ABR48242). Bladder cancer-associated sequences from the present invention
CC
have cytostatic activities, and can be used in antisense gene therapy and
CC
in vaccine production. The method can be used for detecting a bladder
CC
cancer-associated transcript in a cell from a patient. The method is
CC
useful in diagnosing or treating bladder cancer and in screening for
CC
compounds that modulate bladder cancer, such as hormones or antibodies.
CC
The nucleic acid molecules from the present invention may be used in
CC
various screening and diagnostic methods, and for gene therapy, vaccine
CC
and/or antisense/inhibition applications.
XX
Sequence 6906 BP; 1702 A; 1762 C; 1853 G; 1589 T; 0 other;
SQ
Query Match 66.0%; Score 19.8; DB 25; Length 6906;
Best Local Similarity 91.3%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 6 ATCCAGCGCCTCTGTTTGAATCG 28
|||||
Db 2556 ATCCAGCGCCTCTGTTTGAATCG 2578
|||||
RESULT 11
ABS76520
ID ABS76520 standard; cDNA; 6981 BP.
XX
AC ABS76520;
XX
11-DEC-2002 (first entry)
XX
Query Match 66.0%; Score 19.8; DB 24; Length 6981;
Best Local Similarity 91.3%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 6 ATCCAGCGCCTCTGTTTGAATCG 28
|||||
Db 2556 ATCCAGCGCCTCTGTTTGAATCG 2578
|||||
cdNA encoding human ovarian cancer marker OV59.
Human; ovarian cancer; marker; cancer; familial history; brain disorder;
central nervous system disorder; bacterial meningitis; viral meningitis;
Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
brain herniation; inflammation; encephalitis; testicular disorder;
nontuberculous granulomatous orchitis; connective tissue disorder;
heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX
Homo sapiens.
XX
WO200271928-A2.
XX
19-SEP-2002.
XX
14-MAR-2002; 2002WO-US07826.
XX
14-MAR-2001; 2001US-276025P.
XX
14-MAR-2001; 2001US-276026P.
XX
10-AUG-2001; 2001US-311732P.
XX
19-SEP-2001; 2001US-323580P.
XX
26-SEP-2001; 2001US-324967P.
XX
26-SEP-2001; 2001US-325102P.
XX
26-SEP-2001; 2001US-325149P.
XX
(MILL-) MILLENNIUM PHARM INC.
XX
Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
WPI; 2002-723277/78.
XX
P-PSDB; ASG96421.
XX
Assessing whether a patient is afflicted with ovarian cancer, useful in
PT
assessing the stage or progression of the disease, comprises comparing
PT
the expression level of a cancer marker in a sample from a patient and
PT
from a non cancer patient
PT
Disclosure; Page 412-414; 481pp; English.
XX
The present invention relates to a new method for assessing whether a
CC
patient is afflicted with ovarian cancer. The method involves comparing
CC
the expression level of a marker in a patient sample and the normal level
CC
of expression of the marker in a control non-ovarian cancer sample, where
CC
the marker is selected from 363 cancer markers described in the
CC
specification. The method of the invention is useful in diagnosing or
CC
characterising cancer, in detecting the presence of cancer as early as
CC
possible, and the recurrence of ovarian cancer. The method may also be of
CC
particular use with patients having an enhanced risk of developing
CC
ovarian cancer (e.g. patients having a familial history of ovarian
CC
cancer). The cancer markers may be used in the management and treatment
CC
of e.g. brain and central nervous system disorders (e.g. bacterial and
CC
viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC
disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC
inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC
testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC
connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC
disease or atherosclerosis). The compositions and methods may also be
CC
used in assessing the histological type of neoplasm associated with
CC
ovarian cancer, monitoring the progression of ovarian cancer,
CC
determining whether ovarian cancer has metastasized or is likely to
CC
metastasize, selecting a composition for inhibiting ovarian cancer,
CC
assessing the ovarian carcinogenic potential of a compound, or
CC
inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC
present nucleic acid sequence encodes one of the ovarian cancer markers
CC
described in the invention.
XX
Sequence 6981 BP; 1720 A; 1771 C; 1866 G; 1624 T; 0 other;
SQ
Query Match 66.0%; Score 19.8; DB 24; Length 6981;
```

Best Local Similarity 91.3%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGG 28
Db 2631 ATCCAGCGCCTCTGTTTGAATGG 2653

RESULT 12

ABL68557
ID ABL68557 standard; DNA; 6981 BP.

XX AC ABL68557;

XX DT 15-MAY-2002 (first entry)

XX DE 18-JUN-2002 (first entry)

XX DE Human benign prostatic hyperplasia gene #636.

XX KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX OS Homo sapiens.

XX PN WO200212440-A2.

XX PD 14-FEB-2002.

XX PF 07-AUG-2001; 2001WO-US24708.

XX PR 07-AUG-2000; 2000US-223323P.

XX PR 05-JUN-2001; 2001US-0873319.

XX PA (GENE-) GENE LOGIC INC.

XX PA (NISB) JAPAN TOBACCO INC.

XX PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;

XX DR WPI; 2002-257476/30.

XX PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX PT detecting expression levels of one or more genes in prostate cells from
XX PT patient that are differentially regulated compared to normal prostate
XX PT cells

PS Disclosure; Page 344-346; 444pp; English.

XX CC The invention relates to a method of diagnosing (I) the onset or
XX CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
XX CC or identifying an agent that modulates the onset or progression of BPH.
XX CC The method is based on changes in gene expression in BPH tissue isolated
XX CC from patients exhibiting different clinical states of prostate
XX CC hyperplasia as compared to normal prostate tissue. (I) comprises
XX CC detecting the expression levels of one or more genes in prostate cells
XX CC from the subject that are differentially regulated compared to normal
XX CC prostate cells. (II) comprises preparing a first gene expression profile
XX CC of BPH cells or BPH-like cell population, exposing the cells to the
XX CC agent, preparing a second gene expression profile of the agent exposed
XX CC cells, and comparing the first and second gene expression profiles.
XX CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
XX CC useful for identifying an agent that modulates the onset or progression
XX CC of BPH. The methods are useful to present information identifying
XX CC the expression level in a tissue or cells, by comparing the expression
XX CC level of genes given in the specification in the tissue or cells to the
XX CC level of expression of gene in the database, and displaying the
XX CC expression levels of at least one gene in the tissue or cell sample
XX CC compared to the expression level in BPH. Agents using (II) are useful for
XX CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
XX CC benign prostatic hyperplasia gene sequences of the invention.

SQ Sequence 6981 BP; 1720 A; 1771 C; 1866 G; 1624 T; 0 other;

Query Match

Best Local Similarity 66.08; Score 19.8; DB 24; Length 6981;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGG 28
Db 2631 ATCCAGCGCCTCTGTTTGAATGG 2653

RESULT 13

ABL68557
ID ABL68557 standard; DNA; 6981 BP.

XX AC ABL68557;

XX DT 15-MAY-2002 (first entry)

XX DE Kidney cancer related gene sequence SEQ ID NO:6894.

XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX KW stomach; lung; prostate; carcinomas; pancreas; carcinoma; antitumour; cancerous;
XX KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX KW gene; ds.

XX OS Homo sapiens.

XX PN WO200194629-A2.

XX PD 13-DEC-2001.

XX PF 30-MAY-2001; 2001WO-US10838.

XX PR 05-JUN-2000; 2000US-209473P.

XX PR 05-JUN-2000; 2000US-209531P.

XX PR 18-SEP-2000; 2000US-233133P.

XX PR 18-SEP-2000; 2000US-233617P.

XX PR 20-SEP-2000; 2000US-234009P.

XX PR 20-SEP-2000; 2000US-234034P.

XX PR 22-SEP-2000; 2000US-234052P.

XX PR 22-SEP-2000; 2000US-234509P.

XX PR 25-SEP-2000; 2000US-234923P.

XX PR 25-SEP-2000; 2000US-234924P.

XX PR 25-SEP-2000; 2000US-235077P.

XX PR 25-SEP-2000; 2000US-235082P.

XX PR 25-SEP-2000; 2000US-235113P.

XX PR 26-SEP-2000; 2000US-235637P.

XX PR 26-SEP-2000; 2000US-235638P.

XX PR 27-SEP-2000; 2000US-235711P.

XX PR 27-SEP-2000; 2000US-235720P.

XX PR 27-SEP-2000; 2000US-235840P.

XX PR 28-SEP-2000; 2000US-235863P.

XX PR 28-SEP-2000; 2000US-236028P.

XX PR 28-SEP-2000; 2000US-236032P.

XX PR 28-SEP-2000; 2000US-236033P.

XX PR 28-SEP-2000; 2000US-236034P.

XX PR 28-SEP-2000; 2000US-236109P.

XX PR 29-SEP-2000; 2000US-236111P.

XX PR 29-SEP-2000; 2000US-236842P.

XX PR 29-SEP-2000; 2000US-236891P.

XX PR 02-OCT-2000; 2000US-237172P.

XX PR 02-OCT-2000; 2000US-237173P.

XX PR 02-OCT-2000; 2000US-237278P.

XX PR 02-OCT-2000; 2000US-237294P.

XX PR 02-OCT-2000; 2000US-237295P.

XX PR 02-OCT-2000; 2000US-237316P.

XX PR 03-OCT-2000; 2000US-237425P.

XX PR 03-OCT-2000; 2000US-237598P.

XX PR 03-OCT-2000; 2000US-237604P.

XX PR 03-OCT-2000; 2000US-237606P.

XX PR 03-OCT-2000; 2000US-237608P.

XX PR 01-NOV-2000; 2000US-244876P.

XX PR 01-NOV-2000; 2000US-245084P.

XX PA (AVAL-) AVALON PHARM.

XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 XX PI Soppet DR, Weaver Z;
 XX DR WPI; 2002-188264/24.
 XX DR Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX PS Claim 1; SEQ ID 6894; 44pp; English.
 XX CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX SQ Sequence 6981 BP; 1720 A; 1771 C; 1866 G; 1624 T; 0 other;
 Query Match 66.0%; Score 19.8; DB 24; Length 6981;
 Best Local Similarity 91.3%; Pred. No. 53;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 6 ATCCAGCGCCTCTGTTTGATGG 28
 DB 2631 ATCCAGCGCCTCTGTTTGATGG 2653
 RESULT 14
 ABT31940
 ID ABT31940 standard; DNA; 7460 BP.
 AC ABT31940;
 XX 01-MAY-2003 (first entry)
 XX Human breast cancer / ovarian cancer related coding sequence #47.
 XX Human; gene; ds; cytostatic; breast cancer; ovarian cancer.
 XX Homo sapiens.
 XX WO2003000012-A2.
 XX 03-JAN-2003.
 XX 21-JUN-2002; 2002WO-US19773.
 XX 21-JUN-2001; 2001US-300159P.
 XX 27-JUN-2001; 2001US-301351P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Veiby OP;
 XX WPI; 2003-267848/26.
 DR P-PSDB; ABJ37071.
 XX Determining the presence of breast cancer in an individual, involves
 PT using specific polynucleotide markers -

XX PS Disclosure; Page 202-204; 233pp; English.
 XX CC The invention comprises a method for assessing whether a patient is
 CC afflicted with breast cancer or ovarian cancer. The method involves the
 CC use of specific DNA markers. The method of the invention is useful in the
 CC detection and treatment of ovarian and breast cancer. DNA sequences
 CC ABT31894 - ABT31949 encode human breast/ovarian cancer-related proteins.
 XX SQ Sequence 7460 BP; 1844 A; 1890 C; 1992 G; 1734 T; 0 other;
 Query Match 66.0%; Score 19.8; DB 25; Length 7460;
 Best Local Similarity 91.3%; Pred. No. 54;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 6 ATCCAGCGCCTCTGTTTGATGG 28
 DB 3067 ATCCAGCGCCTCTGTTTGATGG 3089
 RESULT 15
 ABX70968/C
 ID ABX70968 standard; cDNA; 570 BP.
 XX AC ABX70968;
 XX 05-MAR-2003 (first entry)
 XX Novel human cDNA sequence #193.
 XX Human; gene; ss; nervous system disorder; peripheral neuropathy;
 KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
 KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
 KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
 KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
 KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
 KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
 KW coagulation disorder; cancer; tumour; inflammatory disease;
 KW septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;
 KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;
 KW haemostatic; antiinflammatory; expressed sequence tag; EST.
 XX Homo sapiens.
 XX WO200281731-A2.
 XX 17-OCT-2002.
 XX 29-JAN-2002; 2002WO-US01222.
 XX 30-JAN-2001; 2001US-0774528.
 XX (HYSE-) HYSEQ INC.
 XX (GOOD/) GOODRICH R W.
 XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Ren F;
 XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2003-058563/05.
 XX Novel polypeptide useful for treating neurodegenerative diseases,
 XX myeloid or lymphoid cell disorders, bone disorders, mechanical and
 XX traumatic disorders, coagulation disorders, and inflammatory diseases
 XX Claim 1; Page -; 612pp; English.
 XX This invention relates to the cDNA sequences encoding an isolated
 CC novel human polypeptide. The protein encoded by the nucleic acid of
 CC the invention is useful for treating central and peripheral nervous
 CC system diseases (e.g. peripheral neuropathy, Huntington's disease,
 CC amyotrophic lateral sclerosis); neurodegenerative diseases (e.g.
 CC Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:41:09 ; Search time 2.7859 Seconds
(without alignments)
4753.042 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tcggatccagcgcctctgtttgatgct 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A-COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B-COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A-COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B-COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PTUS-COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	3	US-09-094-435-1
2	28	93.3	1133	3	US-09-094-435-3
3	19.8	66.0	6642	1	US-08-727-034-5
4	19.8	66.0	6843	1	US-08-727-034-6
5	18.2	60.7	6639	1	US-08-727-034-1
6	18.2	60.7	6961	1	US-08-727-034-2
7	18	60.0	277	4	US-09-313-294A-3050
8	17.8	59.3	33	2	US-08-319-795-14
9	17.8	59.3	33	2	US-08-468-985-14
10	17.8	59.3	1180	4	US-09-220-132-169
11	17.8	59.3	2230	3	US-08-448-194-7
12	17.8	59.3	2230	3	US-08-867-921-7
13	17.6	58.7	360	5	PCT-US93-05703-1
14	17.6	58.7	2262	2	US-08-674-887A-5
15	17.6	58.7	2262	3	US-08-951-844-5
16	17.6	58.7	2262	4	US-09-412-347-5
17	17.6	58.7	2469	1	US-07-997-133-2
18	17.6	58.7	2469	1	US-08-459-296-1
19	17.6	58.7	2469	5	US-07-997-133-2
20	17.6	58.7	2662	2	US-08-451-822A-14
21	17.6	58.7	2662	4	US-08-323-430-14
22	17.6	58.7	2733	1	US-08-371-001-14
23	17.6	58.7	2733	5	PCT-US96-00331-14
24	17.6	58.7	4249	1	US-08-480-784-21
25	17.6	58.7	4249	1	US-08-483-553-21
26	17.6	58.7	4249	1	US-08-487-002-21
27	17.6	58.7	4249	1	US-08-483-554B-21

C 28	17.6	58.7	4249	1	US-08-488-011B-21	Sequence 21, Appl
C 29	17.6	58.7	4249	3	US-08-850-727-21	Sequence 21, Appl
C 30	17.6	58.7	4249	5	PCT-US95-10202-21	Sequence 21, Appl
C 31	17.6	58.7	4249	5	PCT-US95-10203-21	Sequence 21, Appl
C 32	17.6	58.7	4249	5	PCT-US95-10220-21	Sequence 21, Appl
C 33	17.6	58.7	5656	1	US-08-425-061-1	Sequence 1, Appl
C 34	17.6	58.7	5656	1	US-08-825-886-1	Sequence 1, Appl
C 35	17.6	58.7	5656	4	US-08-989-890-1	Sequence 3, Appl
C 36	17.6	58.7	5689	1	US-08-425-061-3	Sequence 3, Appl
C 37	17.6	58.7	5689	1	US-08-825-886-3	Sequence 3, Appl
C 38	17.6	58.7	5689	4	US-08-989-890-3	Sequence 11, Appl
C 39	17.6	58.7	5707	1	US-08-425-061-11	Sequence 11, Appl
C 40	17.6	58.7	5707	1	US-08-825-886-11	Sequence 11, Appl
C 41	17.6	58.7	5707	4	US-08-989-890-11	Sequence 2, Appl
C 42	17.6	58.7	5709	1	US-08-425-061-2	Sequence 7, Appl
C 43	17.6	58.7	5709	1	US-08-425-061-7	Sequence 8, Appl
C 44	17.6	58.7	5709	1	US-08-425-061-8	Sequence 9, Appl
C 45	17.6	58.7	5709	1	US-08-425-061-9	

ALIGNMENTS

RESULT 1
US-09-094-435-1
; Patent No. 6251388
; Sequence 1, Application US/09094435
; GENERAL INFORMATION:
; APPLICANT: Donald L Durden
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
; TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN
; TITLE OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA: US/09/094,435
APPLICATION NUMBER: US/09/094,435
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/049,085
FILING DATE: June 9, 1997
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-094-435-1 Query Match 100.0% Score 30; DB 3; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGATCCAGCGCTCTGTTTGTATGGCT 30
Db 1 TCCGATCCAGCGCTCTGTTTGTATGGCT 30

RESULT 2

US-09-094-435-3
; Sequence 3, Application US/09094435
; Patent No. 6251388

GENERAL INFORMATION:

APPLICANT: Donald L Durden
TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES
TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN
TITLE OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE:

storage

COMPUTER:

IBM Compatible

OPERATING SYSTEM:

IBM P.C. DOS 5.0

SOFTWARE:

FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/094,435

FILING DATE:

Filed Herewith

CLASSIFICATION:

Prior Application Data:

APPLICATION NUMBER: 60/049,085

FILING DATE: June 9, 1997

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/274

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1133 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-094-435-3

Query Match 93.3%; Score 28; DB 3; Length 1133;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCTCTGTTTGTATGGCT 30
Db 96 CGGATCCAGCGCTCTGTTTGTATGGCT 123

RESULT 3

US-08-727-034-5

; Sequence 5, Application US/08727034

; Patent No. 5665872

GENERAL INFORMATION:

APPLICANT: SAITO, YASHUSHI

APPLICANT: IWASAKI, AKIO

APPLICANT: ARAI, KOICHI
APPLICANT: YAMAZAKI, HIROYUKI
TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
TITLE OF INVENTION: THE GENE CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/727,034

FILING DATE: 08-OCT-1996

CLASSIFICATION: 536

Prior Application Data:

APPLICATION NUMBER: JP 261440/1995

FILING DATE: 09-OCT-1995

Prior Application Data:

APPLICATION NUMBER: JP 102451/1996

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 80-079-0

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 6642 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

US-08-727-034-5

Query Match 66.0%; Score 19.8; DB 1; Length 6642;
Best Local Similarity 91.3%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCTCTGTTTGTATGG 28
Db 2434 ATCCAGCGCTCTGTTTGTATGG 2456

RESULT 4

US-08-727-034-6

; Sequence 6, Application US/08727034

; Patent No. 5665872

GENERAL INFORMATION:

APPLICANT: SAITO, YASHUSHI

APPLICANT: IWASAKI, AKIO

APPLICANT: ARAI, KOICHI

APPLICANT: YAMAZAKI, HIROYUKI

TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND

TITLE OF INVENTION: THE GENE CODING THEREFOR

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/727,034
;; FILING DATE: 08-OCT-1996
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 261440/1995
;; FILING DATE: 09-OCT-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 102451/1996
;; FILING DATE: 24-APR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 80-079-0
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6843 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna to mRNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 81..6725
;; OTHER INFORMATION: /note= "Identification Method: S"
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 81..164
;; OTHER INFORMATION: /note= "Identification Method: S"
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 165..6722
;; OTHER INFORMATION: /function= "Nucleotides 165-6722
;; OTHER INFORMATION: encode the mature peptide"
;; OTHER INFORMATION: /note= "Identification Method: S"
;; US-08-727-034-6

Query Match 66.0%; Score 19.8; DB 1; Length 6843;
Best Local Similarity 91.3%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGATGG 28
Db 2514 ATCCAGCGCCTCTGTTTGATGG 2536

RESULT 5
US-08-727-034-1
; Sequence 1, Application US/08727034
; Patent No. 5665872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; TITLE OF INVENTION: THE GENE CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

;;
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/727,034
;; FILING DATE: 08-OCT-1996
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 261440/1995
;; FILING DATE: 09-OCT-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 102451/1996
;; FILING DATE: 24-APR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 80-079-0
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6639 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna to mRNA
;; US-08-727-034-1

Query Match 60.7%; Score 18.2; DB 1; Length 6639;
Best Local Similarity 87.0%; Pred. No. 61;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGATGG 28
Db 2431 ATCCAGCGCCTCTGTTTGATGG 2453

RESULT 6
US-08-727-034-2
; Sequence 2, Application US/08727034
; Patent No. 5665872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; TITLE OF INVENTION: THE GENE CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,034
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 261440/1995
; FILING DATE: 09-OCT-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 102451/1996
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 80-079-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6961 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 178..6819
OTHER INFORMATION: /note= "Identification Method: S"
FEATURE:
NAME/KEY: sig peptide
LOCATION: 178..261
OTHER INFORMATION: /note= "Identification Method: S"
FEATURE:
NAME/KEY: misc feature
LOCATION: 262..6816
OTHER INFORMATION: /function= "Nucleotides 262-6816
OTHER INFORMATION: encode the mature peptide"
OTHER INFORMATION: /notes= "Identification Method: S"
US-08-727-034-2

Query Match 60.7%; Score 18.2; DB 1; Length 6961;
Best Local Similarity 87.0%; Pred. No. 62;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28
|||||
DB 2608 ATCCAGCGCCTCTGTTTGACGG 2630

RESULT 7

US-09-313-294A-3050
Sequence 3050, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 3050
LENGTH: 277
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700610762H1
NAME/KEY: unsure
LOCATION: 3, 11, 15, 37, 44-45, 47, 52, 68, 73, 77, 102, 134, 165, 171, 220, 224
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3050

Query Match 60.0%; Score 18; DB 4; Length 277;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGG 28

DB 200 TCGGATCGAGTCATCCGTTTGTGG 227
|||||

RESULT 8

US-08-319-795-14
Sequence 14, Application US/08319795
Patent No. 5980909
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: Yother, Janet L.
APPLICANT: McDaniel, Larry S
TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
TITLE OF INVENTION: Protein A
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheomaker and Mattare, Ltd.
STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
STREET: Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/319,795
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,636
FILING DATE: 20-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,896
FILING DATE: 20-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/835,698
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/656,773
FILING DATE: 15-FEB-1991
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0810
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-319-795-14

Query Match 59.3%; Score 17.8; DB 2; Length 33;
Best Local Similarity 75.9%; Pred. No. 35;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30
|||||
DB 1 CCGGATCCAGCGTCGTATCTTAGGGCT 29
|||||

RESULT 9

US-08-468-985-14
Sequence 14, Application US/08468985
Patent No. 5997882
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: Yother, Janet L.
APPLICANT: McDaniel, Larry S
TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface

ORGANISM: Homo sapiens
US-09-220-132-169

Query Match 59.3%; Score 17.8; DB 4; Length 1180;
Best Local Similarity 75.9%; Pred. No. 67;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCTCTGTTTGTGGCT 30
Db 533 CTGGATCTGGCGCTCTACTTGGGGCT 561

RESULT 11
US-08-448-194-7/c
; Sequence 7, Application US/08448194
; Patent No. 6028049
; GENERAL INFORMATION:
; APPLICANT: JACOBS, Eric
; APPLICANT: LEGRAIN, Michele
; APPLICANT: MAZARIN, Veronique
; APPLICANT: BOUCHON-THIESEN, Bernadette
; APPLICANT: SCHRYVERS, Anthony B.
; APPLICANT: BLOCH, Marie-Aline
; TITLE OF INVENTION: DNA FRAGMENTS CODING FOR THE TRANSFERRIN
; TITLE OF INVENTION: RECEPTOR OF NEISSERIA MENINGITIDIS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,194
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,469
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/078,053
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92 07493
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016100-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2230 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: DNA which encodes Tbp2 subunit of transferrin
; ORGANISM: receptor
; STRAIN: Neisseria meningitidis IM2169
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 60..119
; FEATURE:

US-09-937-982-1.rn1

ORGANISM: Homo sapiens
US-09-220-132-169

Query Match 59.3%; Score 17.8; DB 2; Length 33;
Best Local Similarity 75.9%; Pred. No. 35;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCTCTGTTTGTGGCT 30
Db 1 CCGGATCCAGCGCTCTGTTTGTGGCT 29

RESULT 10
US-09-220-132-169
; Sequence 169, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Stryjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 1180
; TYPE: DNA

Figure 1

NOTATION: /

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 02:34:35 ; Search time 22.3388 Seconds
(without alignments)
3081.078 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tccggtccagcgcctctgtttgttgcttgcgt 30

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	10	US-09-773-260-1
2	28	93.3	1133	10	US-09-773-260-3
3	19.8	66.0	6868	11	US-09-919-039-39
4	19.8	66.0	6981	12	US-09-873-319-636
5	19.8	66.0	6981	12	US-09-960-706-977
6	19.8	66.0	6981	14	US-10-097-340-299
7	19.8	66.0	7460	10	US-10-176-847-93
8	19.4	64.7	73308	10	US-09-954-456-276
9	19.4	64.7	3186778	13	US-10-027-632-174961
10	18.6	62.7	1010	13	US-10-027-632-255747
11	18.6	62.0	759	13	US-10-027-632-18853
12	18.4	61.3	374	13	US-10-027-632-83004
13	18.4	61.3	378	13	US-10-027-632-38490
14	18.4	61.3	378	13	US-10-027-632-38491
15	18.4	61.3	378	13	US-10-027-632-83005
16	18.4	61.3	574	10	US-09-764-847-62

Sequence 62, Appl	14	US-10-092-154-62	574	14	US-10-092-154-62
Sequence 6300, Ap	9	US-09-815-242-6300	1365	9	US-09-815-242-6300
Sequence 1095, Ap	2407	US-09-764-847-1095	61.3	2407	US-09-764-847-1095
Sequence 2, Appl1	2407	US-10-092-154-1095	61.3	2407	US-10-092-154-1095
Sequence 168244,	12588	US-09-992-009-2	61.3	12588	US-09-992-009-2
Sequence 251412,	832	US-10-027-632-251412	60.7	832	US-10-027-632-251412
Sequence 159, App	1173	US-10-027-632-159	60.7	1173	US-10-027-632-159
Sequence 5159, Ap	2621	US-10-128-714-5159	60.7	2621	US-10-128-714-5159
Sequence 21292, A	488	US-09-918-995-21292	60.0	488	US-09-918-995-21292
Sequence 173748,	854	US-10-027-632-173748	60.0	854	US-10-027-632-173748
Sequence 172306,	865	US-10-027-632-172306	60.0	865	US-10-027-632-172306
Sequence 172307,	1333	US-10-027-632-172307	60.0	1333	US-10-027-632-172307
Sequence 3904, Ap	1686	US-09-764-877-3904	60.0	1686	US-09-764-877-3904
Sequence 3904, Ap	1695	US-09-938-842A-2040	60.0	1695	US-09-938-842A-2040
Sequence 367, App	248	US-09-764-864-367	59.3	248	US-09-764-864-367
Sequence 609, App	281	US-09-998-598-609	59.3	281	US-09-998-598-609
Sequence 7003, Ap	395	US-09-960-352-7003	59.3	395	US-09-960-352-7003
Sequence 12400, A	399	US-09-960-352-12400	59.3	399	US-09-960-352-12400
Sequence 55, Appl	443	US-09-905-243-55	59.3	443	US-09-905-243-55
Sequence 5084, Ap	447	US-09-960-352-5084	59.3	447	US-09-960-352-5084
Sequence 829, App	447	US-09-815-343-829	59.3	447	US-09-815-343-829
Sequence 7936, Ap	477	US-09-960-352-7936	59.3	477	US-09-960-352-7936
Sequence 289, App	544	US-09-998-598-289	59.3	544	US-09-998-598-289
Sequence 388, App	544	US-09-920-300A-388	59.3	544	US-09-920-300A-388
Sequence 388, App	550	US-10-033-528-388	59.3	550	US-10-033-528-388
Sequence 185375,	585	US-10-027-632-185375	59.3	585	US-10-027-632-185375
Sequence 226737,	585	US-10-027-632-226737	59.3	585	US-10-027-632-226737
Sequence 226738,	585	US-10-027-632-226738	59.3	585	US-10-027-632-226738

ALIGNMENTS

RESULT 1
US-09-773-260-1
; Sequence 1, Application US/09773260
; Patent No. US20020102251A1

; GENERAL INFORMATION:

; APPLICANT: Donald L Durden

; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
ASPARAGINASE IN THE TREATMENT OF HUMAN
HEMATOLOGIC AND AUTOIMMUNE DISEASE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSeq for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/773,260

; FILING DATE: 31-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/094,435

; FILING DATE: 1998-06-09

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 234/274

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-773-260-1

Query Match 100.0%; Score 30; DB 10; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30

Db 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30

RESULT 2

US-09-773-260-3

Sequence 3, Application US/09773260

Patent No. US20020102251A1

GENERAL INFORMATION:

APPLICANT: Donald L Durden

TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES

ASPARAGINASE IN THE TREATMENT OF HUMAN

HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: California

ZIP: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/773,260

FILING DATE: 31-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/094,435

FILING DATE: 1998-06-09

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/274

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1133 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-773-260-3

Query Match

Best Local Similarity 93.3%; Score 28; DB 10; Length 1133;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGATGGCT 30

Db 96 CGGATCCAGCGCCTCTGTTTGATGGCT 123

RESULT 3

US-09-919-039-39

Sequence 39, Application US/09919039

Publication No. US20030108871A1

GENERAL INFORMATION:

APPLICANT: Kaser, Matthew R.

TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US

CURRENT APPLICATION NUMBER: US/09/919,039

CURRENT FILING DATE: 2002-09-09

PRIOR APPLICATION NUMBER: 60/222,113

PRIOR FILING DATE: 2000-07-28

NUMBER OF SEQ ID NOS: 401

SOFTWARE: PERL Program

SEQ ID NO 39

LENGTH: 6868

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030108871A1 1269631CB1

US-09-919-039-39

Query Match 66.0%; Score 19.8; DB 11; Length 6868;

Best Local Similarity 91.3%; Pred. No. 25;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2514 ATCCAGCGCCTCTGTTTGATGG 2536

RESULT 4

US-09-873-319-636

Sequence 636, Application US/09873319A

Publication No. US20030134324A1

GENERAL INFORMATION:

APPLICANT: Munger, William E.

APPLICANT: Kulkarni, Prakash

APPLICANT: Getzenberg, Robert H.

APPLICANT: Waga, Iwao

APPLICANT: Yamamoto, Jun

TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic

Hyperplasia Using Gene Expression Profiles

FILE REFERENCE: 44921-5029-US

CURRENT APPLICATION NUMBER: US/09/873,319A

CURRENT FILING DATE: 2001-06-05

EARLIER APPLICATION NUMBER: US 60/223,323

EARLIER FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 755

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 636

LENGTH: 6981

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20030134324A1 U60975

US-09-873-319-636

Query Match

Best Local Similarity 66.0%; Score 19.8; DB 12; Length 6981;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2631 ATCCAGCGCCTCTGTTTGATGG 2653

```

RESULT 5
US-09-960-706-977
; Sequence 977, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 977
; LENGTH: 6981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U60975
US-09-960-706-977

Query Match          66.0%; Score 19.8; DB 12; Length 6981;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||
Db 2631 ATCCAGCGCCTCTGTTTGTATGG 2653

RESULT 6
US-10-097-340-299
; Sequence 299, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen ZHANG
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26

```

```

; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 299
; LENGTH: 6981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-299

Query Match          66.0%; Score 19.8; DB 14; Length 6981;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||
Db 2631 ATCCAGCGCCTCTGTTTGTATGG 2653

RESULT 7
US-10-176-847-93
; Sequence 93, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 7460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-93

Query Match          66.0%; Score 19.8; DB 14; Length 7460;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||
Db 3067 ATCCAGCGCCTCTGTTTGTATGG 3089

RESULT 8
US-09-954-456-2276/c
; Sequence 2276, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27

```

;; PRIOR APPLICATION NUMBER: US/60/235,720
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,840
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,863
;; PRIOR FILING DATE: 2000-09-27
;; NUMBER OF SEQ ID NOS: 2276
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2276
;; LENGTH: 73308
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-954-456-2276

Query Match 64.7%; Score 19.4; DB 10; Length 73308;
Best Local Similarity 79.3%; Pred. No. 48;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCTCTGTTTGTATGGCT 30
Db 19314 CTGGATCCTAGGCTCTTTTGTGGCT 19286

RESULT 9

US-10-027-632-174961
;; Sequence 174961, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108927.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 174961
;; LENGTH: 3186778
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(3186778)
;; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961

Query Match 64.7%; Score 19.4; DB 13; Length 3186778;
Best Local Similarity 79.3%; Pred. No. 64;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCTCTGTTTGTATGGC 29
Db 1767595 TCCCGATCTAGGCTCTCTTTTGACGTC 1767623

RESULT 10

US-10-027-632-255747/c
;; Sequence 255747, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 255747
;; LENGTH: 1010
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-255747

Query Match 62.7%; Score 18.8; DB 13; Length 1010;
Best Local Similarity 76.7%; Pred. No. 60;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCTCTGTTTGTATGGCT 30
Db 447 TCCGGTTCAGGCGCTCTGATGAGCTGGT 418

RESULT 11

US-10-027-632-18853
;; Sequence 18853, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 18853
;; LENGTH: 759
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-18853

Query Match 62.0%; Score 18.6; DB 13; Length 759;
Best Local Similarity 77.8%; Pred. No. 72;
Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GGATCCAGCGCTCTGTTTGTATGGCT 30

```
Db 248 GGCYCCAGCAGCTCTCTGTTGGCT 274
||:||||| ||||| ||||| |||||
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38490

Query Match 61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCGATCCAGCGCCTCTGTTTGTATGCG 29
||| ||||| ||||| ||||| |||||
Db 348 CTGGATCCAGAACTCTGTTTAGTGCG 321

RESULT 14
US-10-027-632-38491/c
; Sequence 38491, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38491
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38491

Query Match 61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCGATCCAGCGCCTCTGTTTGTATGCG 29
||| ||||| ||||| ||||| |||||
Db 348 CTGGATCCAGAACTCTGTTTAGTGCG 321

RESULT 15
US-10-027-632-83005/c
; Sequence 83005, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38490
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83004

Query Match 61.3%; Score 18.4; DB 13; Length 374;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCGATCCAGCGCCTCTGTTTGTATGCG 29
||| ||||| ||||| ||||| |||||
Db 344 CTGGATCCAGAACTCTGTTTAGTGCG 317

RESULT 13
US-10-027-632-38490/c
; Sequence 38490, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38490
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83004
```

```

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83005
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83005

```

```

Query Match      61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY      2 CCGGATCCAGCGCCTCTGTTTGGTGGC 29
Db      348 CTGGATCCAGAAATCTCTGTTTAGGTGGC 321

```

Search completed: August 27, 2003, 11:17:53
Job time : 28.3388 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:35:29 ; Search time 92.0378 Seconds

(without alignments)
7922.115 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tccgaccagcgcctctgtttgatggct 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.6	75.3	789	28	BH249624 BOGAQ68TR
C 2	21	70.0	197	12	BP016090 BP016090
3	21	70.0	530	11	CNS09M4Z
C 4	21	70.0	596	13	BW279822 BW279822

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 5	21	70.0	601	9	AV960848
C 6	21	70.0	644	9	AV674160
C 7	21	70.0	661	9	AV991875
C 8	21	70.0	699	13	BW246571
C 9	21	70.0	758	13	BW246571
10	20.6	68.7	699	9	AU234665
11	20.6	68.7	924	13	BQ931065
12	20	66.7	454	28	AZ557854
13	20	66.7	455	10	BF469784
14	19.8	66.0	416	14	D59983
C 15	19.8	66.0	300	14	C14671
16	19.8	66.0	623	9	AL042811
17	19.8	66.0	882	9	AU132200
18	19.6	65.3	528	9	AA823562
C 19	19.6	65.3	589	9	AI907267
C 20	19.6	65.3	786	28	BH651954
C 21	19.6	65.3	888	13	BU595186
C 22	19.6	65.3	1592	13	BH246061
23	19.6	65.3	4293	11	AK044094
C 24	19.4	64.7	198	10	BG220871
25	19.4	64.7	336	9	AW356036
26	19.4	64.7	345	9	AW427901
27	19.4	64.7	378	9	AA416116
28	19.4	64.7	435	10	BF601890
29	19.4	64.7	468	9	AA590627
30	19.4	64.7	468	9	AA162206
C 31	19.4	64.7	582	10	BE074228
C 32	19.4	64.7	614	28	AZ915332
33	19.4	64.7	618	28	AQ397708
C 34	19.4	64.7	637	29	AG130724
35	19.4	64.7	786	10	BF698692
36	19.4	64.7	896	14	CB198817
C 37	19.4	64.7	914	10	BE789678
C 38	19.4	64.7	976	29	CNS028NY
39	19.4	64.7	1046	13	EX428164
40	19.4	64.7	1175	9	AL515901
41	19.2	64.0	332	14	CA677807
C 42	19.2	64.0	586	29	BZ887403
C 43	19.2	64.0	892	10	BF133459
C 44	19.2	64.0	925	29	CNS033AH
C 45	19.2	64.0	932	29	CNS020SK

ALIGNMENTS

RESULT 1

BH249624

LOCUS

DEFINITION

BOGAQ68TR BOGA Brassica oleracea genomic clone BOGAQ68, genomic

survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH249624 789 bp DNA linear GSS 26-NOV-2001
BOGAQ68TR BOGA Brassica oleracea genomic clone BOGAQ68, genomic
survey sequence.

BH249624.1 GI:17072437

BH249624

BH249624.1

GI:17072437

GSS.

Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 789)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSSs: BOGAQ68TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: Sheared ends.


```

FEATURES
  source
    Location/Qualifiers
      1..789
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"
        /strain="T01000DH3"
        /db_xref="taxon:3712"
        /clone="BOGAQ88"
        /clone_lib="BOGA"
        /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      251 a 130 c 156 g 252 t
ORIGIN
Query Match      75.3%; Score 22.6; DB 28; Length 789;
Best Local Similarity 86.2%; Pred. No. 42;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CCGGATCCAGCGCTCTGTTTGATGGCT 30
Db 512 CCGGAGCCACCGCTCTGTTTGTGCT 540

RESULT 2
BP016090/c
LOCUS
DEFINITION      BP016090 Nori Satoh unpublished cDNA library, young adult Ciona
ACCESSION
VERSION          BP016090.1 GI:19507567
KEYWORDS
SOURCE
ORGANISM
  Ciona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Cionidae; Ciona.
REFERENCE
  1 (bases 1 to 197)
  Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  Expressed genes in Ciona intestinalis
  Unpublished
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: sato@acidian.zool.kyoto-u.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..197
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciad59e6"
        /tissue_type="whole animal"
        /dev_stage="young adult"
        /clone_lib="Nori Satoh unpublished cDNA library, young adult"
BASE COUNT      69 a 45 c 45 g 37 t 1 others
ORIGIN
Query Match      70.0%; Score 21; DB 12; Length 197;
Best Local Similarity 82.8%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CCGGATCCAGCGCTCTGTTTGATGGCT 30
Db 74 CTGGTCCAGGGCTCTGTTTCATGGCT 46

RESULT 3
CNS09M4Z
LOCUS
DEFINITION      Single read from an extremity of a full-length cDNA clone made from Anopheles gambiae total adult females. 3-PRIME end of clone FK0AC48BD03 of strain 6-9 of Anopheles gambiae (African malaria
CNS09M4Z
530 bp mRNA linear HTC 08-JAN-2003
Query Match      70.0%; Score 21; DB 13; Length 596;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

FEATURES
  source
    Location/Qualifiers
      1..596
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciad012b14"
        /tissue_type="gonad"
        /clone_lib="Nori Satoh unpublished cDNA library, gonad"
BASE COUNT      200 a 103 c 89 g 201 t 3 others
ORIGIN
Query Match      70.0%; Score 21; DB 13; Length 596;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

mosquito).
ACCESSION      BX064767
VERSION        BX064767.1 GI:27638048
KEYWORDS      HTC.
SOURCE        Anopheles gambiae (African malaria mosquito)
ORGANISM      Anopheles gambiae
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
REFERENCE      1 (bases 1 to 530)
  Genoscope.
  Direct Submission
  Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Location/Qualifiers
    1..530
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="6-9"
      /db_xref="taxon:7165"
      /clone="FK0AAC48BD03"
      /plasmid="pME18S-FL"
      /note="end : 3-PRIME"
BASE COUNT      99 a 161 c 162 g 108 t
ORIGIN
Query Match      70.0%; Score 21; DB 11; Length 530;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCCGGATCCAGCGCTCTGTTTGATGGC 29
Db 415 TCCGGATCGAGCACCTCTTTTGACGGC 443

RESULT 4
BW279822/c
LOCUS
DEFINITION      BW279822 Nori Satoh unpublished cDNA library, gonad Ciona
ACCESSION
VERSION        BW279822
KEYWORDS      EST.
SOURCE        BW279822.1 GI:24860433
ORGANISM      Ciona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Cionidae; Ciona.
REFERENCE      1 (bases 1 to 596)
  Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
  Expressed genes in Ciona intestinalis (2002c)
  Unpublished
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: sato@acidian.zool.kyoto-u.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..596
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciad012b14"
        /tissue_type="gonad"
        /clone_lib="Nori Satoh unpublished cDNA library, gonad"
BASE COUNT      200 a 103 c 89 g 201 t 3 others
ORIGIN
Query Match      70.0%; Score 21; DB 13; Length 596;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30
Db 72 CTGGGTCCAGGCTTCTGTTTTCATGGCT 44

RESULT 5
AV960848/c
LOCUS AV960848 601 bp mRNA linear EST 14-MAR-2002
DEFINITION AV960848 Nori Satoh unpublished cDNA library, cleavage stage embryo
Ciona intestinalis cDNA clone cici14b08 5', mRNA sequence.
ACCESSION AV960848
VERSION AV960848.1 GI:19449147
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 601)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..601
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cici14b08"
/dev_stage="cleavage stage embryo"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage
stage embryo"
BASE COUNT 176 a 129 c 134 g 162 t
ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 601;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30
Db 580 CTGGGTCCAGGCTTCTGTTTTCATGGCT 552

RESULT 6
AV674160/c
LOCUS AV674160 644 bp mRNA linear EST 05-OCT-2000
DEFINITION AV674160 Nori Satoh unpublished cDNA library Ciona intestinalis
cDNA clone citb14j14 5', mRNA sequence.
ACCESSION AV674160
VERSION AV674160.1 GI:10112159
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 644)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..661
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cici157d16"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage
stage embryo"
BASE COUNT 201 a 135 c 143 g 181 t 1 others
ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 661;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30
Db 620 CTGGGTCCAGGCTTCTGTTTTCATGGCT 592

RESULT 8
BW246571/c
LOCUS BW246571 699 bp mRNA linear EST 09-NOV-2002
DEFINITION BW246571 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb07ic16 5', mRNA sequence.
ACCESSION BW246571

```

Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES

source

Location/Qualifiers

1..644

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="citb14j14"

/tissue_type="whole animal"

/dev_stage="tailbud"

/clone_lib="Nori Satoh unpublished cDNA library"

BASE COUNT 228 a 114 c 90 g 212 t

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 644;

Best Local Similarity 82.8%; Pred. No. 1.9e+02;

Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30

Db 110 CTGGGTCCAGGCTTCTGTTTTCATGGCT 82

RESULT 7

AV991875/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

AV991875 661 bp mRNA linear EST 14-MAR-2002
 AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
 Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.

```

VERSION      BW246571.1  GI:24826489
KEYWORDS     EST.
SOURCE       Ciona intestinalis
ORGANISM     Ciona intestinalis
             Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
             Phlebobranchia; Clonidae; Ciona.
REFERENCE    1 (bases 1 to 699)
AUTHORS      Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE        Expressed genes in Ciona intestinalis (2002c)
JOURNAL      Unpublished
COMMENT      Contact: Nori Satoh
             Department of Zoology
             Kyoto University
             Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
             Tel: 81-75-753-4081
             Fax: 81-75-705-1113
             Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES     source
             1..699
             /organism="Ciona intestinalis"
             /mol_type="mRNA"
             /db_xref="taxon:7719"
             /clone="citb071c16"
             /tissue_type="whole animal"
             /dev_stage="tailbud embryo"
             /clone_lib="Nori Satoh unpublished cDNA library, tailbud
             embryo"

BASE COUNT   211 a 145 c 158 g 185 t
ORIGIN
Query Match 70.0%; Score 21; DB 13; Length 699;
Best Local Similarity 82.8%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTCATGGCT 30
Db 636 CTGGTCCAGGGCTCTGTTTCATGGCT 608

RESULT 9
LOCUS      BU344351
DEFINITION BU344351 758 bp mRNA linear EST 28-NOV-2002
ACCESSION BU344351
VERSION    BU344351.1 GI:25852352
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
             Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 758)
AUTHORS    Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
             Fong,W.R., Tickie,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE      A Comprehensive Collection of Chicken cDNAs
JOURNAL    Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE    22335534
PUBMED     12445392
COMMENT    Contact: Simon Hubbard
             Department of Biomolecular Sciences
             University of Manchester Institute of Science and Technology (UMIST)
             )
             PO Box 88, Manchester, M60 1QD, UK
             Tel: 01612008930
             Fax: 01612360409
             Email: Simon.Hubbard@umist.ac.uk.

FEATURES     source
             1..758
             /organism="Gallus gallus"
             /mol_type="mRNA"
             /strain="Compton Line 151"
             /db_xref="taxon:9031"

```

```

/clone="ChEST467e11"
/sex="Female"
/tissue_type="cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN68"
/notes="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with EcoRI
size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT   190 a 198 c 194 g 176 t
ORIGIN
Query Match 70.0%; Score 21; DB 13; Length 758;
Best Local Similarity 82.8%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCCGATCCAGCGCCTCTGTTTCATGGC 29
Db 647 TCAGCATCCAGCGCGCTGTTTCATGGGC 675

RESULT 10
LOCUS      AU234665
DEFINITION AU234665 699 bp mRNA linear EST 26-OCT-2001
ACCESSION AU234665
VERSION    AU234665.1 GI:16500812
KEYWORDS   EST.
SOURCE     Branchiostoma belcheri
ORGANISM   Branchiostoma belcheri
             Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
             Branchiostoma.
REFERENCE  1 (bases 1 to 699)
AUTHORS    Suzuki,M.M. and Satoh,N.
TITLE      Genes expressed in the amphioxus notochord revealed by EST analysis
JOURNAL    Dev. Biol. 224 (2), 168-177 (2000)
MEDLINE    20387132
PUBMED     10926757
COMMENT    Contact: Nori Satoh
             Department of Zoology
             Kyoto University
             Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
             Tel: 81-75-753-4081
             Fax: 81-75-705-1113
             Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES     source
             1..699
             /organism="Branchiostoma belcheri"
             /mol_type="mRNA"
             /db_xref="taxon:7741"
             /clone="Bb5_02D02"
             /tissue_type="notochord"
             /clone_lib="Amphioxus Notochord cDNA library"

BASE COUNT   181 a 159 c 211 g 148 t
ORIGIN
Query Match 68.7%; Score 20.6; DB 9; Length 699;
Best Local Similarity 85.2%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 GGATCCAGCGCCTCTGTTTCATGGCT 30

```

TITLE
Mouse BAC End Sequences from Library RPCI-23
and Flase1/C.M.
Mouse BAC End Sequences from Library RPCI-23

cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..456
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-aut-d-03-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_M_S4"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_M_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged, normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

BASE COUNT 82 a 132 c 95 g 146 t 1 others

Query Match 66.7%; Score 20; DB 10; Length 456;
Best Local Similarity 82.1%; Pred. No. 4.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGC 29
|||||
Db 212 CCGTTCCAGTCCTCTGTTTGATGCC 239
|||||

RESULT 14

D59983
LOCUS HUM079C11B Clontech human fetal brain polyA+ mRNA (#535) Homo sapiens cDNA clone GEN-079C11 5', mRNA sequence.
DEFINITION

ACCESSION D59983.1 GI:961622

VERSION D59983

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 215)

AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.
Fujiwara et al. (1995)

TITLE

JOURNAL COMMENT

Unpublished
Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.
Location/Qualifiers

FEATURES

source
1..215
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GEN-079C11"
/clone_lib="Clontech human fetal brain polyA+ mRNA (#6535)"

BASE COUNT 51 a 54 c 54 g 56 t

ORIGIN

Query Match 66.0%; Score 19.8; DB 14; Length 215;
Best Local Similarity 91.3%; Pred. No. 4.7e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28
|||||

Db 27 ATCCAGCGCCTCTGTTTGATGG 49
|||||

RESULT 15

CI4671/c

LOCUS CI4671 Clontech human aorta polyA+ mRNA (#6572) Homo sapiens cDNA

DEFINITION clone GEN-079C11 3', mRNA sequence.

ACCESSION CI4671

VERSION CI4671.1 GI:1569378

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 300)

AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.
Fujiwara et al. (1995)

TITLE

JOURNAL COMMENT

Unpublished
Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.
Location/Qualifiers

FEATURES

source
1..300
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GEN-079C11"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"

BASE COUNT 70 a 80 c 80 g 67 t

ORIGIN

Query Match 66.0%; Score 19.8; DB 14; Length 300;
Best Local Similarity 91.3%; Pred. No. 5.1e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28
|||||

Db 189 ATCCAGCGCCTCTGTTTGATGG 167
|||||

Search completed: August 27, 2003, 09:30:45
Job time : 98.0378 secs

